Preemption

and the

Counterfactual Analysis

of

Causation

by

David Anthony Coady

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1999

Approval

This manuscript has been read and accepted for the Graduate Faculty in Philosphy in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

May 26 1999
Date

Chair of Examining Committee

Aug 30 1999
Date

Executive Officer

Professor Hartry Field
Professor David Lewis
Professor Arnold Koslow
Professor Michael Levin
Professor Richard Mendelsohn

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Preemption and the Counterfactual Analysis of Causation

by

David Anthony Coady

Adviser: Professor Hartry Field

Since David Lewis first published "Causation" it has been a widely accepted dogma that the most straightforward counterfactual analysis of causation (which I call, "the naive analysis") cannot succeed, because of a class of counterexamples which, following Lewis's nomenclature, have come to be called cases of "preemption". Consequently, there has been a debate amongst philosophers including Tim Maudlin, Paul Horwich, Jonathan Bennett, David Armstrong, Martin Bunzl, Douglas Ehring, Ned Hall, Michael McDermott, Richard Miller, Murali Ramachandran, Laurie Paul, Jonathan Schaffer, and others, about how to respond to this "fact". I argue that this debate is premature, since the naive analysis has not been refuted, or been shown to conflict with any intuitions that deserve to be respected. There is no such thing as preemption.

Acknowledgments

Many people have helped me during this project, and I apologize to those who are not mentioned here. All mistakes are, of course, my own.

I want to thank all the members of my comittee for their help, and especially my supervisor Hartry Field, without whom this would have been very different in both time and manner (*i.e.*, it would have been later and worse) and, of course, David Lewis, without whom not even a very robust candidate for being this dissertation could have occurred at all.

Samir Chopra, Jonathan Schaffer, Martin Bunzl, Mark Johnston, Jared Blank, and Boran Beric, have all given me encouragement and constructive criticism. Tony, Margaret, Phyllis, and Benjamin Coady have been a source of inspiration, as have Joe and Patsy Habeeb.

Finally, I want to dedicate this dissertation to my wife, Leanne Habeeb, whose emotional and intellectual support made it possible.

Table of Contents

Introduction

The guiding idea:

The guiding idea behind the counterfactual analysis of causation is that a causal proposition is a kind of counterfactual proposition.[1]

There are different kinds of causal proposition. I will be following most of the contemporary literature on causation by concentrating on singular causal propositions, *i.e.*, propositions which express a causal relation between particular things, rather than general causal propositions, *i.e.*, propositions which express a

---

[1] The reader may substitute "claim" or "belief" for "proposition" in what follows. I don't believe that anything I say will depend on whether or how these terms should be distinguished; or whether one rather than another are the primary bearers of truth values.

causal relation between kinds of things.[2] I think that there is a sense in which the former are more fundamental. For the moment, however, that must remain a hunch. I think that general causal propositions may also be analyzable in counterfactual terms. That too must remain a hunch for now.

I will also follow the lead of most of the contemporary literature on causation by calling the particular things which stand in causal relations 'events'.[3] I hope to persuade the reader that this does not beg any important philosophical questions about the nature of causation. The reader will not go far wrong, if she understands my use of the term 'events' as a way of referring to whatever particular things stand in causal relations. This inevitably means that I will refer to some items as 'events', which we would not ordinarily think of as such.

For example, we would ordinarily think of events as involving change. Both *my snubbing of my friend*, and *my continuing to exist*,

---

[2] An example of this latter kind of causal proposition is "Smoking causes lung cancer". For discussion of the relation between these two kinds of causal proposition see Cartwright, Eells and Hitchcock.

[3] See Jonathan Bennett *Events and their Names* (1988) for an exception.

should, however, in my view, be counted as events, since they appear to have causes and effects. Like David Lewis, I am unconcerned about how it sounds to call such things 'events'.[4]

For similar reasons I am inclined to treat *omissions* (*i.e.*, the failure of events to occur) as themselves being events: for example *the failure of an explosion to occur* and *the failure of that failure to occur*, should both, in my view, be counted as events in the same sense that explosions are.[5]

If there should turn out to be a good argument that unchanges or omissions do not really stand in causal relations, then I will stop calling them 'events'. Although I uncover no such arguments, I do not rule out the possibility of being persuaded to change my mind; it will make little or no difference to my thesis.[6]

The only remotely substantive claims about events that I am committed to are, firstly, and most importantly for any counterfactual analysis of causation, that events are things which

---

[4] David Lewis, "Events", *Philosophical Papers Vol. 2.*, p. 261. Henceforth, "Events".

[5] In fact the latter of these omissions will be an explosion.

[6] David Lewis argues that omissions are not events (or anything else) in his unpublished "Void and Object".

may meaningfully (though not necessarily elegantly) be supposed not to occur (or happen, or take place). Secondly, they must be localized; that is, they must occur in a specific spatiotemporal region.

The 'Naive' analysis:

I defend what I will call the naive counterfactual analysis of event causation, *i.e.*, the claim that a causal proposition of the form '*c* causes *e*' (where c and e are distinct events) is true iff the counterfactual proposition 'If c had not occurred, then e would not have occurred' is true.

Restricting the scope of the analysis to distinct events is necessary, if we are to exclude certain noncausal relations. Although there is no standard definition of 'distinctness' in the literature, it seems clear that it is at least partly a mereological notion: two events are not distinct, if they have a common part.

Every event counterfactually depends on itself, but few events, if any, cause themselves. It also seems to be the case that without The First World War, The Battle of the Somme would not have occurred, but this does not of course mean that The First World War

caused The Battle of the Somme. Counterfactual dependence sometimes indicates a mereological, rather than a causal, relation between events.

Counterfactual dependence can also indicate a logical rather than a causal relation between events. It seems to be the case, for example, that when one slams a door, one also closes it: without the closing, the slamming would not occur. This does not mean, however, that the closing causes the slamming.[7] While there seems to be some sense in which these events are not distinct, it is not obvious what that sense is. We could say that 'they' are identical, and therefore not distinct but, for reasons that will become clear in Chapter Five, I do not think we can assume this. We could say that they have a common part, but it is not intuitively clear that we are entitled to say so.

In fact, as we shall also see in Chapter Five, it seems to be an implication of David Lewis's theory of events that the closing and the slamming may well have a common part, in virtue of the latter literally being a part of the former. Unfortunately Lewis's theory of

---

[7] The example is from Jaegwon Kim, "Causes and Counterfactuals" *Journal of Philosophy*, 70 (1973) pp. 571.

events has ontological commitments (*i.e.*, unactual entities) that many, myself included, are reluctant to endorse. I think it remains an important task for any adequate counterfactual analysis to articulate a satisfactory sense of 'distinctness'.

Despite its simplicity, and its consistency with the guiding idea that a causal proposition is a kind of counterfactual proposition, I know of no contemporary author who is prepared to defend the naive counterfactual analysis. The standard objections to it have nothing to do with any lack of clarity about the notion of *distinctness*, however.

The author who comes closest to my position is David Hume. For two reasons I disagree with Hume about the form of the relevant counterfactuals, however, which he identifies as 'if the first object had not been, the second never had existed'.[8] Firstly, this unnecessarily restricts the analysis to causal propositions which assert that the existence of one object caused the existence of another. I think that such propositions are best seen as a proper subset of propositions which assert that the occurrence of one event caused the occurrence of another. Secondly, I suspect that specifying

---

[8] *An Enquiry Concerning Human Understanding*, Section VII

the temporal order of the events (or objects), as Hume did, is undesirable and unecessary.

Are there cases of causation in which the usual counterfactuals don't hold?

While everyone, except for me, agrees that the naive counterfactual analysis is a nonstarter, defenders of some kind of counterfactual analysis (as well as those who reject the counterfactual approach altogether) are divided amongst themselves over whether the biconditional fails in both directions, *i.e.*, the truth of the causal proposition is neither necessary nor sufficient for the truth of the counterfactual proposition (as Tim Maudlin argues), or just one direction, *i.e.*, the truth of the causal proposition is sufficient but not necessary for the truth of the counterfactual proposition (as David Lewis argues).

David Lewis and many of those influenced by his work would agree that the counterfactual proposition entails the causal proposition, but claim that there are realistic thought experiments about *redundant causation* which show that the entailment doesn't go

in the other direction. Without exception the extensive literature which has been generated from Lewis's discussion of these problem cases accepts his claim that they provide counterexamples to the naive analysis. Consequently disagreement has been about how to modify the naive analysis, while still retaining something which deserves to be called a 'counterfactual analysis of causation'. I will argue that this debate is premature, to say the least, because the naive analysis has not been refuted.

Tim Maudlin shares the widespread view that the truth of a causal proposition is not sufficient for the truth of its corresponding counterfactual proposition. He offers the following example which purports to show the possibility of causation without the usual counterfactual dependence:

> The electrical activity in the primary navigational computer was among the causes of the shuttle's perfect landing, for example, but even had that activity not occurred the shuttle would still have landed in the same way: the back up computer would have kicked in.[9]

---

[9] Tim Maudlin, *Quantum Non-Locality and Relativity* p.127. Henceforth, Maudlin (1994).

It has become standard in the literature, following Lewis's nomenclature, to call cases like this *preemption;* I will discuss in detail a range of such cases.

As an advocate of the naive analysis I cannot afford to admit that there are any genuine cases of preemption, since they would constitute counterexamples to that analysis. I will attempt to make the denial of preemption plausible by arguing that each putative example of it should be redescribed in one of two ways; which way is best will usually depend on context.

In some contexts, I will argue, the alleged preemption is in fact *symmetrical overdetermination*; in such contexts none of the redundant causes count as genuine causes (or at least they are not genuinely separate causes). In other contexts, I will argue, the alleged *preempting cause* is in fact a staightforward (*i.e.*, nonredundant) cause.

Are there cases in which the relevant counterfactual holds, but the causal proposition doesn't?

Maudlin, Bennett, and Horwich have all argued that the analysis fails in the other direction as well, *i.e.*, that the truth of the causal proposition is not necessary for the truth of the corresponding counterfactual proposition. There are, they argue, circumstances in which it would be correct to say that if c had not occurred, neither would e, but wrong to say that c caused e, even though c and e are distinct events.[10] Maudlin argues that this could be the case if c was an effect of e rather than a cause of it:

> Suppose, for example, that one is playing pool, trying to sink the 7 ball in the corner by a banked shot. Further suppose that in order to make the shot the ball must carom off a particular small section of the cushion. If the shot is made then the ball hitting that area of cushion is causally implicated with it falling into the pocket: if it hadn't bounced off that particular area it would not have gone into the pocket. But it is also arguable that the counterfactual dependence goes the other way around as well: if it hadn't gone into the pocket then it couldn't have hit the right section of the cushion.[11]

---

[10] Strictly speaking Lewis (or at least a temporal part of him) would not disagree. However he would claim that such counterfactuals are true only under a nonstandard interpretation which is irrelevant to the analysis of causation.

[11] Maudlin (1994) pp. 128-129

In fact Maudlin thinks that it is more than just arguable that this cause counterfactually depends on this effect; he thinks that it is true; although he confines his argument for this position to a footnote.[12]

The issue is not strictly whether 'backtracking counterfactuals' are true.[13] It would not be credible to suppose that they are never true (or at least assertable). Rather, the issue is whether they are true in the same sense as their converses are. Maudlin agrees with Horwich and Bennett that they are, Lewis argues that they are not.

Maudlin claims that this is not the only way in which it can be true that if c had not occurred, then neither would e, without c causing e, and even though *c* and *e* are distinct. In fact he thinks that a counterfactual of this form could be true without either of the relevant events being a cause or an effect of the other. This could happen if c and e are effects of a common cause:

---

[12] ibid.

[13] For clarity I am giving 'backtracking counterfactual' a somewhat narrower meaning than Lewis, from whom I take the terminology, does. A backtracker, for me, is a counterfactual which asserts that if a certain event had not occurred, then neither would one of its causes.

> Immediately after the satisfying thunk of the 7 ball falling in the corner pocket, the cue ball drops dishearteningly into the side. In fact, any shot which manages to sink the 7 also scratches. So if the cue ball hadn't scratched, the 7 wouldn't have gone in either. Or so, at least, the expert might inform the amateur who laments "I would have won with that shot if only the cue ball hadn't gone into the side."[14]

Again, the issue is not strictly over whether 'effects-of-a-common-cause counterfactuals' can be true. There are undeniably contexts in which they do seem to be true (or at least assertable). Rather the issue is whether they are true in the same sense in which 'causal counterfactuals' are true. Maudlin agrees with Bennett and Horwich that they are, Lewis argues that they are not.

Maudlin draws the conclusion that counterfactuals should be used, not to analyse causation itself, but the weaker relation of *causal connection*. Two events x and y are causally connected if x causes y, y causes x, or x and y are effects of a common cause.

I do not think there is a need to restrict our ambition in this way, since Lewis is right that *backtracking* and *effects of a common cause* counterfactuals are not true in the sense in which causal counterfactuals are. It is not necessary to claim, as Lewis has, that

---

[14] ibid. p.129

causal counterfactuals are 'standard' whereas the others are 'nonstandard'; it will be enough to argue that the causal ones are different from the other two, and that this difference can be explicated without appeal to causal notions.

A full defense of this position, however, would require a reasonably detailed account of the truth conditions of counterfactuals, which allowed us to distinguish those underlying the *is a cause of* relation on the one hand, from those which underlie both the *is an effect of* relation and the *is an effect of a common cause of* relation, on the other hand, and to do so without circularity. Although I am confident that this can be done, it is beyond the scope of this dissertation.

Summary:

The approach taken by Lewiseans to alleged cases of *preemption* has been to add various bells and whistles to the analysis to deal with those cases in which it is alleged that a causal

claim is true, even though its corresponding counterfactual is not.[15] These bells and whistles are, I will argue, unnecessary. Furthermore, in several cases in which an analysis with one or the other of these bells or whistles comes into conflict with the 'naive' analysis it is the latter which delivers the more plausible result.

This is not to say that the naive counterfactual analysis will please everyone. In some rather far-fetched thought experiments it leads to counterintuitive consequences. However, I argue that it would be a mistake to conclude from this that the naive analysis fails to capture the meaning of 'causation'; the counterinuitive consequences can be explained away.

---

[15] I borrow David Lewis's terminology for such modifications of the naive analysis.

Chapter One

Preemption and the Counterfactual Analysis

Paul Horwich has claimed that the pendulum has swung in favour of regularity, as opposed to counterfactual, analyses of causation.[16] I don't believe this is true. To a large extent they appear to suffer from shared difficulties.[17] The two accounts both have a problem dealing with preemption, but, as we shall see, it represents a different kind of problem for each account. It seems to me that the pendulum should swing toward whichever account deals best with its problem of preemption, and I think that a counterfactual analysis can handle its problem without too much pain. It can do so by denying that it exists.

For Lewis *preemption* is a kind of *redundant* causation. There are in turn different kinds of redundant causation, but Lewis restricts his discussion to the following case:

---

[16] Paul Horwich, *Asymmetries in Time*, (Massachusetts, 1989) p.176. Henceforth, Horwich (1989)

[17] Arguably they both have difficulty distinguishing genuine causes from effects and epiphenomena.

> Suppose we have two events c_1 and c_2, and another event e distinct from both of them; and in actuality all three occur; and if either one of c_1 and c_2 had occurred without the other, then also e would have occurred; but if neither c_1 nor c_2 had occurred, then e would not have occurred. Then I shall say that c_1 and c_2 are *redundant causes* of e.[18]

The naive counterfactual analysis will count neither c_1 nor c_2 as causes of e, since e does not depend counterfactually on either of them. It may, however, count the combination of c_1 and c_2 (assuming there is such an event) as a cause of e.

Lewis divides redundant causation into preemption and (*symmetrical*) *overdetermination*.[19] In a case of symmetrical overdetermination c_1 and c_2 have an equal claim to being causes of e. In such cases, it is typically unclear whether c_1 and c_2 are each causes of e, or whether neither are, but it is clear that there is no reason to say that one is, whereas the other is not, a cause of e.

---

[18] David Lewis, "Causation: Postcript E", *Philosophical Papers vol. 2* (Oxford, 1986) p.193. Henceforth, "Postscript E". Lewis takes this to be redundant causation in its simplest form. He passes over redundant causation in other forms. In particular he passes over 'probabilistic redundant causation'.

[19] In what follows it will occasionally be convenient to refer to preemption as 'asymmetrical redundancy' and (symmetrical) overdetermination as 'symmetrical redundancy.

Because it is unclear what to say about such cases Lewis correctly treats them as being poor test cases for analyses of causation. This is not true of cases of preemption however:

> In a case of preemption, the redundant causes are not on a par. It seems clear that one of them, the *preempting cause*, does the causing; while the other, the *preempted alternative*, waits in reserve. The alternative is not a cause; though it could and would have been one, if it had not been preempted.[20]

As Lewis notes, whereas counterfactual analyses of causation have difficulty saying why the preempting cause *is* a cause, regularity accounts have difficulty saying why the preempted alternative *is not* a cause.[21] The problem for the counterfactual account is to say why the preempting cause is a cause when the effect is not dependent on it.

So what, according to Lewis, is the difference between the *preempting cause* and the *preempted alternative* in virtue of which the former but not the latter is a cause of the event *e*?

---

[20] ibid. p.199

[21] Roughly speaking, the problem for the regularity analysis is to explain why the preempted alternative is not a cause in those cases in which it is just as highly correlated with the effect as the preempting cause is.

In answering this question a division should be made between what I will call 'Early Lewis' and 'Late Lewis'. Although, as we shall see, the published discussions of preemption have been largely prompted by Lewis's later views, I will argue that in one important respect first thoughts were best.

Although I will be arguing that there is no such thing as preemption, as defined by Lewis, it will sometimes be convenient to use the term "preemption" to refer to putative examples of it.

Chapter Two

Early Lewis

Early Preemption, Transitivity, and Causal Chains

But I do not think that Lewis's first thoughts were exactly right; nor could I afford to, given my commitment to what I have called the naive counterfactual analysis of causation.

Lewis's first published counterfactual analysis of causation (*i.e.*, Early Lewis) is closer to the simple one that I am defending, than that of Late Lewis, however there is one feature that distinguishes it from my account. Lewis originally thought that this difference could solve the problem of preemption.[22] Counterfactuals are used to analyse a relation he calls *causal dependence*. Causation, in turn, is analysed in terms of the ancestral of this relation:

> Let c, d, e, ... be a finite sequence of actual particular events such that d depends causally on c, e on d, and so on throughout. Then this sequence is a *causal chain*.

---

[22] As we shall see Lewis came to think that this modification alone could not do the work.

> Finally, one event is a cause of another iff there exists a causal chain leading from the first to the second.[23]

Apart from its alleged help with preemption, Lewis (and many other advocates of a counterfactual analysis of causation) argue that this addition of theoretical complexity is required to make causation transitive. It is a common assumption that causation is transitive, and as a result both critics and advocates of some form of counterfactual analysis of causation have tended to take for granted that such an analysis must include (at least) this modification in order to be worthy of serious discussion. But it is arguable that the alleged transitivity of causation is vulnerable to counterexamples. I borrow the following apparent counterexample from Hartry Field.

Suppose *c* is my enemy lighting the fuse to the bomb he has previously planted in my apartment, *d* is my friend stamping out the fuse outside my door, and *e* is my drawing breath a moment after the bomb was set to explode.

On Lewis's account *c* causes *e*, because there is a causal chain from *c* to *e*. Intuitively, however, it is the simple account of causation, and not Lewis's account which gets the right result here;

---

[23] "Causation" p. 167.

c causes *d* and *d* causes *e*, but *c* does not cause *e*. By contrast with Lewis's analysis, the naive analysis seems to get the desired result in this case; namely that causation inherits the intransitivity which is an uncontroversial feature of counterfactuals.[24]

Originally Lewis assumed that the only genuine cases of preemption were those in which a branch process diverging from the main process connecting the preempting cause with *e* inhibits the alternate process which would otherwise have connected the preempted alternative to *e*. Following Lewis's terminology I will call this *early preemption*. I offer the following example of it.

Suppose the preempting cause is Charley's decision to kill Edgar, the preempted alternative is Cecil's decision to kill Edgar, *e* is Edgar's death; Cecil sees Charley on his way to kill Edgar and consequently abandons his murderous intention, because he judges it better to let Charley kill Edgar. In this case the branch process is the light reflecting from Charley to Cecil's eyes; it inhibits a

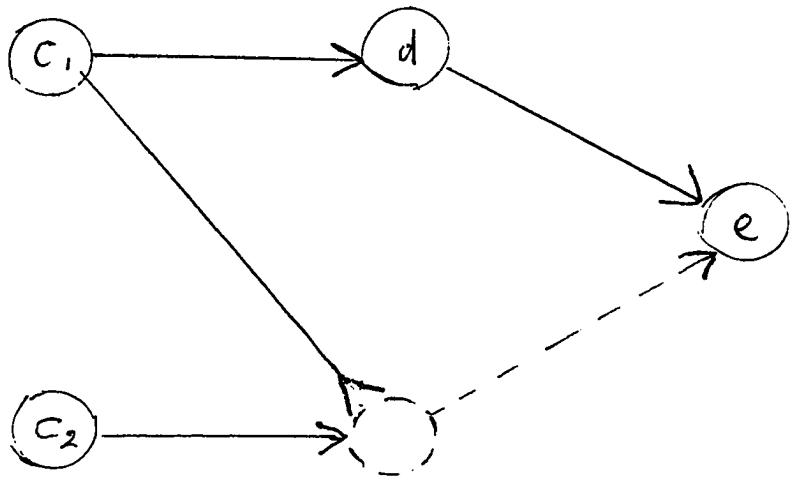
---

[24] For other apparent counterexamples to the transitivity of causation see Michael McDermott, "Redundant Causation" *British Journal for the Philosophy of Science* (46, 1995) pp. 531-533. McDermott does not comment on this, but the naive analysis seems to get the right result for each of these examples, as it does for Field's example.

process initiated by Cecil which, if not inhibited, would have led Cecil (perhaps redundantly) to kill Edgar.

This kind of situation can also be illustrated by the following 'neuron diagram':

Diagram 1:



In this system there are stimulatory synapses (shown by the forward arrowheads) and an inhibitory synapse (shown by the backward arrowhead). A neuron fires if stimulated, but not if it is inhibited at the same time. Neurons c_1 and c_2 fire; setting off two processes which make their way toward neuron *e*. The preempting process which begins with the firing of c_1 goes to completion. The preempted alternative is prevented from doing so by a branch

process that diverges from the main process at a junction event before it can stimulate neuron *e*.

Neuron diagrams like this have become, since Lewis's work, a popular way of representing causal interactions. There is a potentially misleading feature of such Lewisean neuron diagrams, which requires a brief digression. I have been following Lewis's practice of using the symbols c_1, c_2, and *e* to refer to events. In these diagrams, however, the same symbols are used to refer to neurons, which are certainly not events.[25] Lewis is, of course, aware of the distinction between a neuron and the event of its firing. I will argue, however, that a tendency in the literature to conflate what ought to be kept distinct in this respect partly explains why preemption has seemed to be such an intractable problem. This is one way in which the 'problem of preemption' is, I hope to show, attributable to ways of representing (both linguistic and diagrammatic) putative examples of preemption.

What is the difference between the firing of c_1 and the firing of c_2 in virtue of which the former, but not the latter, causes *e*?

---

[25] Lewis uses large case letters to refer to neurons. I shall not follow this practice.

> As far as causal dependence goes, there is no difference: e depends neither on c_1 nor on c_2. If either one had not occurred, the other would have sufficed to cause e. So the difference must be that, thanks to c_1, there is no causal chain from c_2 to e; whereas there is a causal chain of two or more steps from c_1 to e. Assume for simplicity that two steps are enough. Then e depends causally on some intermediate event d, and d in turn depends on c_1.[26]

According to Early Lewis there must be some intermediate event along the preempting process which can be used to form a two-step causal chain from c_1 to e. Any event along the preempting process will do, providing it is between the neuron where the branching process begins (from now on the branching neuron) and the neuron where the actual and the unactualized processes merge (from now on the merging neuron).

It will sometimes be convenient to refer to the firing of the branching neuron as 'the branching event' and the firing of the merging neuron as 'the merging event'. In Diagram 1, like most published neuron diagrams illustrating early preemption, the effect is the merging event. This is not essential: in Lewis's illustration of this phenomenon the firing of the merging neuron causes the effect.

---

[26] "Causation" pp.171-172.

This will make no difference to my discussion, until Chapter Seven.[27]

Objection 1:

If *d* had not occurred, then *e* would still have ocurred, since if *d* had not occurred then neither would its cause c_1, so c_2 would not have been preempted from causing *e*.

Reply:

This objection is anticipated by Lewis. He warns us that this would be 'backtracking`. Originally Lewis had argued that counterfactuals like 'If *d* had not occurred then neither would *e*' are not true.[28] He later argues that there is a sense in which they are true, but that this sense is 'nonstandard'. Counterfactuals which are true only in a nonstandard way are, for Lewis, out of place in tracing causal dependence.[29]

---

[27] It is another nonessential feature of my diagram that the preempting cause is also the branching event: the preempting cause may instead be, as it is in Lewis's diagram, an event which causes the branching event. This will make no difference to my discussion.

[28] "Causation" pp. 170-171

[29] David Lewis "Counterfactual Dependence and Time's Arrow" *Philosophical Papers vol. 2*, (Oxford, 1986). *This* distinction between Lewis's earlier and later views is unrelated to the distinction

A detailed discussion of this putative distinction between standard and nonstandard counterfactuals is beyond the scope of this dissertation; Michael McDermott, however, has suggested that we don't have to engage in such a discussion, to accept Lewis's response to this objection. We only have to accept that a counterfactual like 'If *d* had not occurred, then neither would *e*' has "two reasonable interpretations":

> ... one on which it is true, and one on which it is false. We can then stipulate that the reading for the counterfactuals used in the analysis of causal statements is to be the former.[30]

But we can not rest content with this partial defence of Lewis, since it is central to many of the published criticisms of Lewis's account that his postulation of an ambiguity between two kinds of counterfactuals is mistaken.[31]

Since, as I argued in the introduction, synonymy between backtracking and causal counterfactuals would be as fatal for my

---

between what I have called 'Early Lewis' and 'Late Lewis'.

[30] Michael McDermott, "Redundant Causation", p.524

[31] See Jonathan Bennett, "Counterfactuals and Temporal Direction" *Philosophical Review* (93, 1984) pp.57-91 and Horwich (1989) pp. 161-162.

analysis as it would for Lewis's, I am committed to the view that a genuine ambiguity does exist, and can be explained without appeal to unanalysed causal notions. Unfortunately I cannot defend this position here.

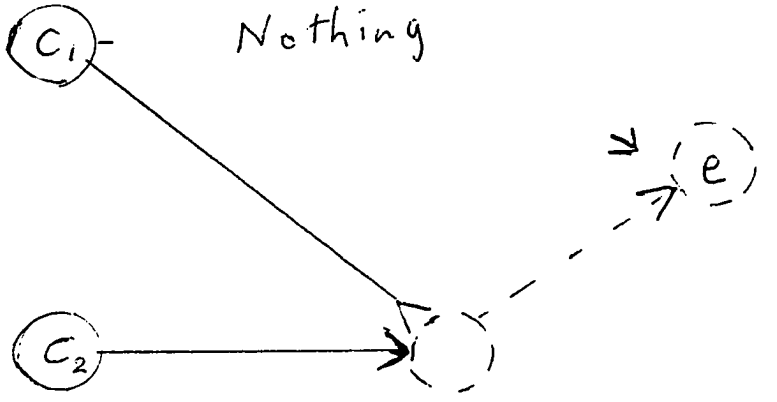
Objection 2:

Why assume that there is an intermediate event along the main process between the branching neuron and the merging neuron? Why, that is, can't the branching event cause the merging event directly?[32]

Reply:

Lewis anticipates this possibility, which he illustrates with another diagram:

Diagram 2:



[32] This objection is made by Horwich (1989) pp.171-172, and Martin Bunzl "Causal Preemption and Counterfactuals" *Philosophical Studies* (37, 1980) p.117.

Although Lewis classified this a case of *late* preemption, for reasons that will become clear in the next chapter, it is more useful to think of it as a kind of *early* preemption.[33] He thought that such cases are possible, and he concedes that his account does disagree with commonsense; which seems to make c_1 a genuine cause, and c_2 a merely potential cause, just as in the previous diagram.

Since such cases involve action at a distance, however, which goes against "what we take to be the ways of this world", he claims that agreement with commonsense in such cases is "not an urgent goal".[34]

Lewis came to think that another kind of preemption, genuinely late preemption, which he had originally thought unproblematic, needed a quite different treatment. I will argue that his first thoughts were right on this topic. The fact that, if late preemption were a problem, it could not be solved by Lewis's introduction of

---

[33] According to my definition this should be considered *early* rather than *late* preemption, since the preempted process is inhibited by a branch process leading from the preempting process, rather than by a continuation of the preempting process beyond the effect.

[34] "Postscript E" p. 203

causal chains, however, strengthens the case against introducing this modification in the first place.

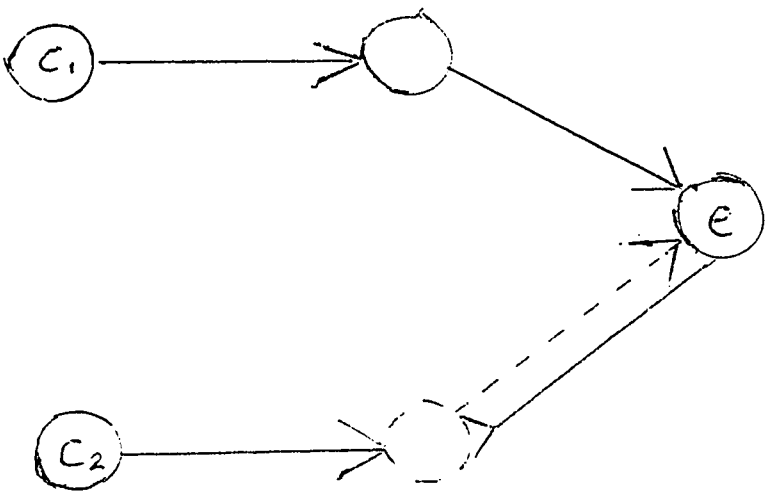
Chapter Three

Late Lewis

Late Preemption, and Quasi-Dependence

In cases of late preemption the preempted process is "cut off, not by a branch process that diverges from the main process at a junction event before the effect is reached, but rather by a continuation of the main process beyond the effect."[35] Suppose this time that Cecil desists only when he sees Edgar dead. Lewis illustrates late preemption with the following system of neurons:

Diagram 3:



[35] ibid. p. 203

In such cases it is quite clear that Early Lewis's modification of the naive analysis does not help. There is no event which can be found along the preempting process which can be used to complete a chain of causal dependence, since the preempted process is not doomed until after the effect has ocurred.[36]

Lewis came to think that these cases presented a problem for his theory. Originally he had dismissed them because there seemed to be a straightforward solution. Early Lewis thought that in such cases the actual firing of *e* can and should be distinguished from the counterfactual firing of *e* which would have resulted from the preempted process. If so, then the counterfactual analysis (both my own and Lewis's) get the desired result that c_1's firing is a cause of the actual firing of *e* and c_2's firing is not.

Before discussing why Late Lewis disagreed with Early Lewis, and why I think that in this respect early Lewis was right, I will

---

[36] I am following Lewis here in discounting the possibility of backward causation and simultaneous causation. This is not because either of us think these things are impossible, but because they are contrary to the way we take our world to be. Such possibilities, like the possibility of action at a distance, may be treated as spoils to the victor.

briefly discuss Late Lewis's approach to the problem of late preemption.

Quasi-Dependence:

Lewis used the notion of *quasi-dependence* to make a second modification to the naive analysis, which he hoped would deal with late preemption. Lewis is aware that it would be outrageous to include such modifications *for the sole purpose* of getting rid of counterexamples. Consequently the introduction of both these modifications is accompanied by a justification, which is intended to be independent of the specific problem it is intended to solve.

The causal chain modification was introduced by Early Lewis to satisfy the alleged intuition that causation is transitive; and incidentally to solve the only kind of preemption problem which Lewis recognized at the time, *i.e.*, the problem of early preemption. Similarly, the quasi-dependence modification was introduced by late Lewis to satisfy the alleged intuition that causation is an intrinsic characteristic of processes (or the regions in which they take place); and incidentally to solve the problem of late preemption.

Lewis claims that it is intuitively clear that whether a process is causal depends only on the intrinsic nature of the process itself, and on the relevant laws of nature. Suppose there is a process, call it P, which is not, in the sense previously defined, a causal chain. Suppose, however, that P is in its intrinsic character just like processes in other regions (of the same world, or other worlds with the same laws). Suppose further that amongst these processes the great majority - as measured by variety of the surroundings - are causal chains. Lewis's original analysis would not count such P as causal, since it is not a causal chain. He came to think, however, that this was a problem for his original analysis, since absence of causation should not be a result of "bad surroundings"; something intuitively irrelevant to the presence or absence of causation:

> So we might extend the analysis. Suppose that there exists some actually occurring process of the kind just described, and that two distinct events *c* and *e* are the first and last in that process. Then let us say that *e quasi-depends* on *c*. We might wish to count that as one kind of causation, based derivatively on counterfactual dependence even though there is no dependence between those two events themselves. As before, we must take an ancestral to ensure that causation will come out transitive; thereby providing not only for chains of stepwise dependence, but also for chains of stepwise

> quasi-dependence, or mixed chains. To this end we could redefine a *causal chain* as a sequence of two or more events, with either dependence or quasi-dependence at each step. And as always, one event is a *cause* of another iff there is a causal chain from one to the other.[37]

Despite its added complexity and its increased distance from the guiding idea of counterfactual analyses of causation, Late Lewis endorsed this approach, because of its help with late preemption and its consistency with the 'intuition of intrinsicness'.

It seems that quasi-dependence can help with the problem of late preemption because the absence of a causal chain (in the old sense) from c_1 to e is a result of something extraneous to the process which connects them; namely the presence alongside the main process of one or more preempted alternatives:

> Without them, all would be well. Hold fixed the laws but change the surroundings, in any of many ways, and we would have the dependence that my original analysis requires for causation. But as it is, we have quasi-dependence instead of dependence. So if we extend the analysis, and allow causation by quasi-dependence, that solves our problem. We then can agree with common sense that we have genuine preemption, and genuine causation by the preempting cause.[38]

---

[37] ibid. p. 206

[38] ibid. pp. 206-207

Late Lewis's appeal to the intuition of intrinsicness and his solution to the problem of late preemption are brought together. Although the main - *i.e.*, preempting - process, in Diagram 3, does not exhibit the required pattern of causal dependence to count as a causal chain (in the old sense) it gets to be counted as a causal chain (in the new sense), because of its intrinsic similarity to other possible processes which are causal chains (in the old sense):

> The complication would be objectionable if it were just a hoky gimmick to deal with late preemption, but it is not just that. For what it is worth, we also have independent motivation in the intuition of intrinsicness.[39]

According to Late Lewis extending the concept of a causal chain in this way vindicates common sense twice over. Firstly it allows us to agree with his intuition that if two processes are intrinsically similar, then either both are causal or neither are. Secondly it allows us to agree with his intuition that the main - *i.e.*, preempting - process is causal, and the alternate - *i.e.*, preempted - process is not.

Objection 1:

---

[39] ibid. p. 207

Why suppose that, in cases of late preemption, the vast majority of nomically possible processes intrinsically similar to the main process will be causal chains (in the old sense)?

The best answer I can think of, on Lewis's behalf, is that we should suppose this because the 'cousins' of the main process will tend to be unlike it in not being accompanied by any late preempted alternatives.[40] However it's not clear - to me at least - that there is any reason to think that processes which are intrinsically just like Cyril's (second) killing of Edgar will tend to occur in the absence of a late preempted alternative. It may seem that there is such a reason; namely the relative scarcity of murder attempts. This answer may help with this example, but there seems no reason to suppose that it will help with all cases of late preemption.[41]

Is there any reason to think that processes which are intrinsically just like the main process of Diagram 3 will tend to occur in the absence of any late preempted alternative? I can't think

---

[40] I borrow the term from Murali Ramachandran, "A Counterfactual Analysis of Causation", *Mind* (106, 1997) p. 265. Processes are cousins iff they have the same intrinsic character and are subject to the same laws.

[41] In fact I think this answer will not help with this example either, but it would require too much of a digression to explain why.

of any. It seems to be nomically possible for such a process to take place without being accompanied by any alternative process. That is, there will be some processes which are intrinsically just like the preempting process (in its world, or other worlds with the same laws) which are not themselves late preempting processes. But this is not to say that the great majority of such processes will have this characteristic.

One possible way to avoid this difficulty would be to weaken Lewis's analysis. Suppose that instead of quasi-dependence requiring the great majority of nomically possible processes to exhibit the proper pattern of dependence, we only require that *some* nomically possible process exhibit this pattern. I'm not sure how Lewis would respond to this suggestion. There seems to be a danger that it would make causation too easy. Rather than dwell on this, however, I want to discuss a possible case of late preemption which is equally (if at all) a problem for either interpretation of quasi-dependence. That is, it is a *prima facie* problem for any attempt to 'piggyback' the causal structure of late preempting processes on the counterfactual structure of intrinsically similar processes.

Objection 2:

Suppose there were an example of late preemption with this feature: the laws of nature guarantee that the preempting cause must be accompanied by its preempted alternative.[42] In such cases the effect will not be quasi-dependent on the preempting cause in either of the senses of 'quasi-dependence' I have considered.

Lewis allows for this posssibility. He insists, however, that even in such cases we should still be able to form a causal chain (in the new extended sense) connecting the preempting cause to the effect. The first step in constructing such a chain is to take some intermediate event *d* along the main process: like all the events, prior to *e*, along the main process *d* will depend on c_1 (*i.e.*, the preempting cause under consideration). Since there is no event along the main process on which the effect is dependent *e* will not depend on *d*, but *e* will quasi-depend of *d*. Thus we have a process from c_1 to *d* to *e* with either dependence or quasi-dependence at each step; satisfying the extended definition of a causal chain. So c_1 is a cause of *e*; which for Lewis is the desired result. So long as we have some

---

[42] Lewis considers this. He credits John Etchemendy with bringing the possibility of such a case to his attention.

event along the main process on which *e* quasi-depends Lewis's analysis can deal with it.

Late Lewis concedes that his account is not immune to counterintuitive results in some possible cases; cases in which there is no intermediate event to do the job. He thinks, however, that such cases would, like the second kind of early preemption, go against our most fundamental beliefs about the kind of world we inhabit:

> What if there is *no* intermediate that could lawfully have been produced without also producing a preempted alternative? That makes the case very peculiar indeed. It is central to the way we ordinarily think about preemption that we can regard the main and the alternative processes as distinct and separable. So if the laws forbid us to have even a part of the one process without the corresponding part of the other, that goes badly against our habitual presuppositions. If so, such common sense opinions as we may have need not be respected - spoils to the victor.[43]

Suppose we put aside whatever concerns we might have about Lewis's cavalier attitude towards far-fetched cases. It is still not clear to me why laws which forbid parts of preempting processes without the corresponding parts of preempted processes should go against our habitual presuppositions any more than laws which

---

[43] ibid. footonote 27

forbid the occurrence of a preempting cause without the occurrence of a preempted alternative.

It may be that Lewis would be better off ignoring both of these possibilities; treating them each as spoils to the victor. This approach would, as Lewis notes, buy him some simplicity elsewhere in his theory, since it would allow him to "drop the part of the analysis in which we take an ancestral to ensure that causation turns out transitive."[44] After all, the quasi-dependence modification seems to solve the problem of early preemption just as well (or just as poorly) as it does the problem of late preemption.

Objection 3:

Murali Ramachandran has pointed out that Lewis's solution seems to have difficulty explaining why the preempted alternative is not a genuine cause.[45] What reason is there for saying that there is a causal chain (in the new sense) from the firing of c_1 to the firing of e which is not also a reason for saying that there is such a chain from the firing of c_2 to the firing of e?

---

[44] ibid.

[45] M. Ramachandran, (1997) p. 265.

Ramachandran offers a reply on behalf of Lewis that seems to be a reasonable interpretation of what Lewis had in mind. Whereas it is nomically possible for processes which are intrinsically indistinguishable from the main process to occur without any preempted alternative; it is nomically impossible for processes which are intrinsically indistinguishable from the preempted alternative to occur without being preempted. The reason for this asymmetry is the incompleteness of the alternative process. Processes which are intrinsically indistinguishable from the alternative process, except for the fact that they are completed, cannot really be intrinsically indistinguishable from it. This is because they include an event, such as the firing of neuron *d*, which has no counterpart in the alternative process. There is no such difference between the main process and its cousins.

This answer will work on the assumption that there will always be an event which did not occur, but would have occurred along the alternative process, had that process not been preempted. It may be that Lewis is entitled to treat cases of late preemption in which this assumption does not hold as spoils to the victor; however

too many cases of this kind suggests that we are looking in the wrong direction.

Objection 4:

Lewis's guiding intuition that intrinsically similar processes are either both causal or neither causal is arguably counterintuitive, and vulnerable to counterexamples.

Michael McDermott claims to have an example of two intrinsically indistinguishable processes occurring in different worlds with the same laws; one of which is causal, the other of which is not:

> World w_1 is pretty much like the actual world. One day, in a fit of bad temper, Nixon reaches for the button. Instantly Haig leaps forward, throws Nixon to the ground, and sits on his head until he calms down. Next morning, Joe Blow eats breakfast as usual. If Haig had not leapt into action, Nixon would have pressed the button, triggering a nuclear attack on Russia; the Russian response would have destroyed all human life in America; Blows breakfast would have remained uneaten. Haig's leap, then, was a cause of Blow's eating breakfast ...[46]

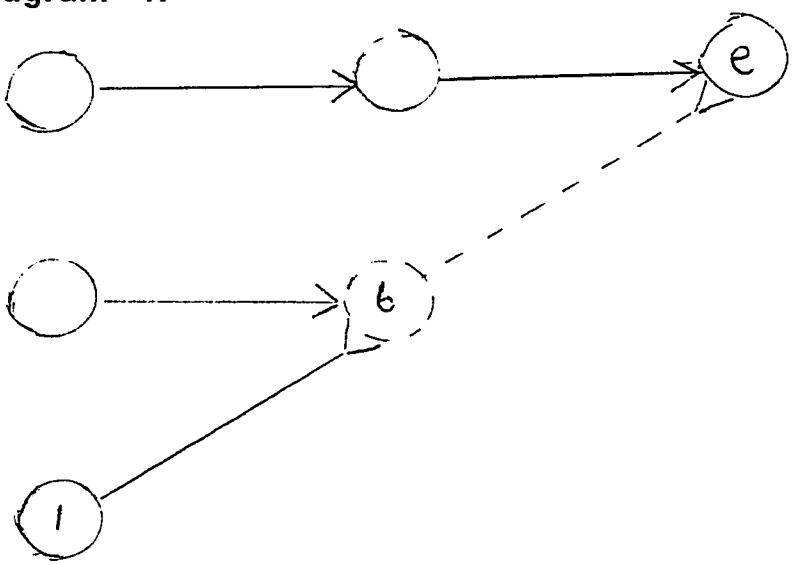
This is a kind of case, which Ned Hall has called 'double prevention', in which a process goes to completion, because its potential aborter

---

[46] "Redundant Causation" p. 529

is itself aborted.[47] Hall illustrates double prevention with the following neuron diagram :[48]

Diagram 4:



In this diagram, the firing of *l* corresponds to Haig's leap, the failure of *b* to fire corresponds to Nixon's failure to press the button, and the firing of *e* corresponds to Blow's eating breakfast. If Haig had not leapt, Blow would not have eaten his breakfast; so, on the counterfactual analysis (both my own and Lewis's), Haig's leap

---

[47] Ned Hall "New Problems for an Analysis of Causation" *unpublished manuscript* p. 19

[48] The letters have been changed to fit McDermott's example. David Armstrong has suggested that calling this "double prevention" is misleading since it suggests two preventions of the main process. He suggests the term "meta-prevention", since it involves the prevention of a prevention. Since "double prevention" has gained such currency in the literature, however, I will continue to use it.

causes Blow to eat his breakfast.[49] McDermott asks us to compare the events of w_1 with those of w_2:

> World w_2 shares the laws of w_1, and the relevant events in America are the same: Haig's leap prevents Nixon's pressing the button, and thereby prevents the launching of the nuclear attack; Blow eats breakfast next morning. But in w_2 Russia is entirely uninhabited. If the American attack had been launched ... there would have been no counterattack; Blow would have eaten his breakfast next morning undisturbed. It seems clear that in w_2 Haig's leap was *not* a cause of Blow's eating breakfast. The series of events <Haig's leaping, Blow's eating> is a causal process in w_1, it occurs in an intrinsically identical region in w_2, but it is not a causal process in w_2[50]

The corresponding neuron diagram for this scenario would be just like that for double prevention, except for the absence of an inhibitory synapse from *b* to *e*.

It would be possible for Lewis to deny that the pairs <Haig's leaping, Blow's eating> and <*I*'s firing, *e*'s firing> are genuine processes, on account of their distance from one another. This would, however, seem to be in conflict with the fact that these

---

[49] Hall denies this, and argues that double prevention is fatal for any counterfactual analysis of causation.

[50] ibid.

events, in the original examples, are one step causal chains (in the old sense).

Rather than pursue questions about the nature of processes, or their intrinsic characteristics any further, I will assume the reader shares my skepticism about the prospects for this approach. Lewis himself no longer thinks this approach was a good idea.[51]

---

[51] Recent graduate seminar in Princeton. This is part of a view emerging as I write, which might be called 'Latest Lewis'.

Chapter Four

The common suggestion

Appealing to the fragility of the effect to dissolve redundancy

A more obvious strategy for saving, not just the spirit, but the letter, of the counterfactual analysis would be to deny that Lewis's examples of early and late preemption do exhibit redundancy.[52] We should agree of course that without the firing of c_1 in these examples, *e* would still have fired. But why not say that this counterfactual firing of *e* would have been a different event from the actual firing of *e*? Likewise we might wonder what is to stop us from supposing that the actual death Edgar suffered at the hands of Charley would have been a different death from the one Cecil would have inflicted on him.

---

[52] The examples I have in mind are those illustrated in Diagrams 1, 2, and 3. If the suggestion under consideration were right, then they would not really be examples of preemption, at least according to Lewis's definition, since that is a kind of redundant causation. I will continue my practice of using the word "preemption" to refer to these examples, for the sake of convenience.

In order to say something sensible about this strategy we need some conception of what it takes, in Lewis's words, to make "a numerical difference between an event that actually occurs and one that would have occurred under some counterfactual supposition."[53] An adequate counterfactual analysis, it appears, needs to be supplemented with acceptable criteria for the trans-world identity of events. Jonathan Bennett puts the need as follows:

> In order to counterfactualize about a particular item x, we must be able to distinguish worlds at which x is present-but-different from worlds that lack x altogether. So, as Lewis makes clear, any counterfactual about an event implies something about its essence.[54]

Peter van Inwagen's view that the causes of events are of their essence would, if acceptable, vindicate the strategy under consideration.[55] If this were right the actual and counterfactual firings of *e* (and deaths of Edgar) must be different events, since they would have been produced by different causes.

---

[53] "Postscript E" p. 195

[54] Jonathan Bennett, *Events and their Names* (Indianapolis, 1988) p. 55. Henceforth, Bennett (1988).

[55] Peter van Inwagen in "Ability and Responsibility," *Philosophical Review* (87, 1978) pp. 208-209, and *An Essay on Free Will* (Oxford, 1983) pp.167-170.

Lewis rejects Inwagen's view for two reasons. Firstly, it would make any attempt to analyse causation in terms of counterfactuals look suspiciously circular: making trivial "any counterfactual to the effect that without the cause, the effect would not have occurred."[56] Secondly, we do appear to be able to suppose that one and the same event might have been caused differently.[57] Lewis rightly recognizes that this appearance may be deceptive. But he makes a good *prima facie* case against Inwagen's doctrine, and I will assume that it is false.

Lewis assumes that the essence of an event is to be found in its time or manner of occurrence.[58] To the extent that an event could not have occurred at a different time or in a different manner, it is, in Lewis's terminology, "a fragile event". A fragile event has

---

[56] "Postscript E" p. 195

[57] ibid.

[58] ibid. I see no reason not to include place on an equal footing with time. I assume that Lewis does not mention the possibility of spatial location being relevant to event identity, only because there is no question of the actual and counterfactual effects in these examples occurring in different places.

"stringent conditions of occurrence"; *i.e.*, it could very easily have failed to happen, because of its "rich" essence.[59]

So how fragile are events? Lewis suggests that this is the wrong question. Instead we should be asking: what are "our standards of fragility"?[60] The difference between these two questions will not matter for the purposes of this chapter, but I hope to show that the suggestion that the essences of events are in some sense, and to some degree, up to us is an important insight, the implications of which have been widely overlooked, perhaps even, to some extent, by Lewis himself.

Like Inwagen's causal essentialism, extreme fragility would allow us to distinguish the actual and counterfactual firings of *e* (and deaths of Edgar):

> It is a common suggestion to adopt extreme standards of fragility, and thereby make away with redundant causation altogether. Even if a man is shot dead by a firing squad, presumably it would have made *some* minute difference to the time and manner of his death if there had been seven bullets instead of eight. So if you fired one of the eight bullets, that made some difference;

---

[59] "Postscript E" p. 196

[60] ibid. pp.196-97

> so if his death is taken to be very fragile indeed, then it would not have occurred without your act.[61]

If we were to adopt this suggestion in order to get rid of redundant causation, we would *ipso facto* get rid of preemption.

We have two dimensions along which we can measure degree of fragility: time and manner of occurrence. This means that strictly speaking there are two kinds of fragility: *temporal* fragility and *manner of occurrence* fragility. As we will see, the former offers the possibility of getting rid of late preemption, the latter offers the possibility of getting rid of early preemption (or at least all realistic cases of it).

If the effect, in a putative case of late preemption, is temporally fragile enough the redundancy (and hence the preemption) disappear. Since, the alternate process is blocked by the effect itself, the counterfactual effect would have taken place after the actual effect.[62] Consequently a sufficiently (temporally) fragile effect will always be distinguishable from the effect which would have resulted from the alternate process.

---

[61] ibid. p.197

[62] Remember that, for reasons already discussed, Lewis and I are ignoring the possibility of backward and simutaneous causation.

We could get rid of the problem of late preemption then by advocating *temporal essentialism*: this is the doctrine that for every event the time at which it occurs is of its essence. Menzies has advocated this way of dismissing late preemption.[63] It is natural to suppose that dissolving late preemption by appealing to the fragility of the effect *requires* the truth of the doctrine of temporal essentialism. How else could we be sure the effect is temporally fragile enough? I hope to show, in the next chapter, that we can get rid of late preemption without being temporal essentialists, and without adopting uniformly stringent standards of temporal fragility.

Why did Lewis come to reject a strategy - adopting temporally fragile standards for the late preempted effect - which seems to solve a difficult problem for him so neatly? Lewis says that it will not work because "fragility of the effect is no better as a remedy for these cases of late preemption than it is as a remedy for redundant causation generally".[64]

---

[63] Peter Menzies, "Probabilistic Causation and Causal Processes: A Critique of Lewis", *Philosophy of Science* (56, 1989) p. 653. Henceforth, Menzies, (1989).

[64] "Postscript E" p.204

Lewis raises two problems with adopting "extreme standards of fragility" to eliminate redundant causation generally. The first of these is that such standards would be inconsistent with the way we talk about events. After all, he observes, we "do say - within limits! - that an event could have been postponed and could have happened differently."[65] He does not consider this decisive, however, since "the standards that apply within the analysis of causation might differ from those which apply in explicit talk".[66]

I think that Lewis is right that we should prefer, all else being equal, a theory which allows that at least some events could have occured at times and in ways other than those they actually do. I also think that Lewis is right that this consideration should not be thought of as decisive.

If the essences of events can not be reliably inferred from our counterfactual suppositions about them, what else can guide us? Jonathan Bennett attributes the following answer to Lewis:

> ... start with our firm beliefs about what causes what, put them into their counterfactual form in accordance with the analysis, and draw conclusions about what the

---

[65] ibid. p.198

[66] ibid.

> essences of events must be like if we are not to be convicted of too much error in our causal beliefs ...[67]

This method gives rise to Lewis's second argument against attributing extremely rich essences to events; an argument which he treats as particularly significant and which has had the most influence in the literature.

Too Many Causes?

For Lewis the main difficulty with extreme fragility is that, in combination with the counterfactual analysis of causation, it creates countless spurious causes for each event:[68]

> Boddie eats a big dinner, and then the poisoned chocolates. Poison taken on a full stomach passes more slowly into the blood, which slightly effects the time and manner of the death. If the death is extremely fragile, then one of its causes is the eating of the dinner. Not so.[69]

---

[67] Bennett, (1988) p. 61. See also Bennett, (1987) p. 369.

[68] Lewis attributes the argument to Ken Kress.

[69] "Postscript E" p. 198

The problem with extreme fragility is that it implies that anything which even slightly influences the time or manner of an event counts as a cause of it.[70]

The specific problem with temporal fragility is that it implies that anything which even slightly influences the time of an event counts as a cause of that event:

> It is one thing to postpone an event, another to cancel it. A cause without which it would have occurred later, or sooner, is not a cause without which it would not have occurred at all. Who would dare be a doctor if the hypothesis under consideration were right? You might manage to keep your patient alive until 4:12, when otherwise he would have died at 4:08. You would then have caused his death. For his death was, in fact, his death at 4:12. If that time is essential, his death is an event that would not have occurred had he died at 4:08, as he would have done without your action. That will not do.[71]

---

[70] This argument against extreme fragility has been extremely popular. Peter Menzies, (1989) p. 649, argues against extreme fragility by citing the undesirability of treating an explosion at the center of the sun, which produces a neuron, which in turn passes through a person's body as she dies, as a cause of death. Jonathan Bennett, (1988) p. 65, cites the unacceptability of supposing that Plato's wiping of Socrates' brow caused his death in support of the same conclusion.

[71] "Events" p. 250

I assume the reader shares Lewis's (and my own) reluctance to say that the *doctor's treatment* and *Boddie's dinner eating* are causes of their respective deaths. This does not show, however, that they are not causes. Sometimes we are reluctant to say things because they are false; sometimes we are reluctant to say things for purely pragmatic reasons. We could agree with Lewis that it is undesirable to say, for example, that a doctor's delaying of a death by four minutes should count as a cause of that death; while accepting that it is nevertheless the literal truth. Our reluctance to assent to what is the literal truth may be explicable by the fact that it implies (in the everyday sense, not the logical sense) what *is* false; namely that the doctor is ethically and legally responsible for the death.

It is fairly common in the literature on causation to stipulate that the concept of causation under discussion is egalitarian; *i.e.*, it is free of salience conditions. Sometimes when we say that one event caused another event we mean that it was a particularly important cause (or at least that it was important in the context in question). This is what what we seem to be doing when we single out one event as *the* cause (as opposed to *a* cause) of another. Lewis says

that he is not concerned with this kind of restricted concept of causation:

> 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 nondiscriminatory concept of causation.[72]

It seems plausible to suppose that such an analysis will endorse causal propositions which we feel reluctant to accept.

It seemed to Lewis that the fact that temporal essentialism entails that anything which even slightly affects the time of an event counts as a cause of that event is a *reductio* of temporal essentialism. When we remind ourselves, however, that we are dealing with an egalitarian causal relation in which unimportant, irrelevant, and misleading causes are all genuine causes, the unacceptability of this consequence appears to be reduced.

Lewis assumed that his examples show that a slight difference in time need not constitute a numerically distinct event. Suppose this were so. This implies that if the *doctor's treatment* (or the *dinner eating*) had delayed death long enough, then we would be more willing to count it as a cause of death. But this seems to be wrong.

---

[72] "Causation" p. 162, see also Bennett, (1989) pp. 70-71.

Who would dare be a doctor if delaying death by any amount counted as killing. If death were delayed longer we would not be any more likely to count the treatment as a cause of it.

This suggests that it is not the fact that the doctor's treatment only *slightly* influences the time of death that makes us reluctant to say that it caused it: it is the fact that it influenced the time of death by making it later than it would otherwise be.

Suppose we accept the intuition we are asked to have about Lewis's examples; namely that these death delayers are not even pragmatically ignorable causes of death. This would show that the counterfactual analysis is inconsistent with temporal essentialism. If delaying an event does not cause it, then that same event must have been able to occur earlier than it actually did. So at least some events could have occurred earlier than they actually do.

Accepting this possiblility, however, is not inconsistent with treating alleged cases of late preemption as spurious. In order to rule out the possiblility of a particular effect being late preempted, it is enough that we rule out the possibility that it could have occurred later than it actually did.

Let us call the doctrine that no event could have occurred later than it did *one way temporal essentialism.* This position, which is weaker than the temporal essentialism advocated by Menzies, is just as adequate for dissolving late preemption problems. Furthermore it respects the intuitions Lewis appeals to with his examples of slightly delayed deaths.

Jonathan Bennett has argued that a counterfactual analysis of causation cannot allow that an event could have occurred later than it did, without also allowing that it could have occurred earlier than it actually did.[73] I do not agree. Suppose we have an asymmetric accessibility relation. If Lewis is right about the doctor's treatment and Boddie's death there is a possible world *w*, accessible from the actual world, in which a certain death occurs four minutes before it does in the actual world. If, as we are for the moment supposing, deaths (or events in general) could have occurred earlier than they actually did, but not later, the actual world cannot be accessible from *w*. If it were accessible, then the proposition that the death could have occurred later than it did would be true at *w*.

---

[73] Jonathan Bennett, "Event Causation: The Counterfactual Analysis", *Philosophical Perspectives* 1 (1987)

Of course, one way temporal essentialism does not entail that any events could have occurred earlier than they actually did, it merely allows for the possiblility. Furthermore it is not just a hokey gimmick dreamt up to get rid of late preemption problems, but has some independent support in the fact that we tend to count hasteners of events as causes, but not delayers. This fact may partly explains our intuitions about Lewis's examples.

This is not to say that one way temporal essentialism is entirely free of problems. It will have problems with any examples of hasteners which do not seem to be causes. The following example, from Penelope Mackie, is as compelling as any I know of:

> Suppose that I have made plans to spend September in Paris, but the receipt of a subpoena requiring my attendance at the Old Bailey in September leads me to reschedule my trip to August. Receiving the subpeona hastens my visit to Paris, but it would be odd to say that it causes my visit - or even that it is one of the causes of my visit.[74]

If no event could have occurred later than it actually did, then the trip to Paris would not have taken place had the subpeona not arrived. Consequently the arrival of the subpeona counts as a cause

---

[74] Penelope Mackie, "Causing, Delaying, and Hastening: Do Rains Cause Fires?", *Mind* (1992) p. 485.

of the visit, according to one way temporal essentialism, despite the intuition Mackie asks us to share.[75]

We could respond by reminding ourselves again that we are working with a concept of causation which is free of salience conditions. Is it plausible to suppose that the subpoena *is* one of the causes of the visit - just not a very significant one? The plausibility of this position seems to increase when one reflects on the fact that there are contexts in which we would be inclined to count it as a cause. Arguably there are contexts in which we would be inclined to treat each of the time affecting noncauses in the three examples under consideration as genuine causes: we would tend to do this in those contexts in which the time of the candidate effect is particularly salient.

---

[75] Mackie's intuition, if correct, would also be a problem for L. A. Paul's analysis, see L.A. Paul, "Keeping track of the Time: Emending the Counterfactual Analysis of Causation" *Analysis* (59, 1999). Paul exploits our tendency not to count delayers as causes to solve the problem of late preemption. Her approach does not build the temporal asymmetry into the essence of the effect, but into the counterfactual supposition about it. Causal dependence is analysed, not by a counterfactual with the form 'If *c* had not occurred, then *e* would not have occurred', but by one with the form 'If c had not occurred, then e would not have occurred or would have occurred later than it did'. Her analysis, like Lewis's, also differs from the naive analysis in analysing causation in terms of the ancestral of this relation.

Lewis's too-many-causes objection to temporal essentialism, as well as to other forms of what he would consider 'extreme' fragility, work by asking us to agree that not just anything which influences the time or manner of an event is a cause of that event. It is open to temporal essentialists (one way or otherwise), or for that matter anyone seeking to argue that alleged cases of asymmetric redundancy are not genuine cases of redundancy at all, to reply that our reluctance to count these slight affecters as genuine causes, is solely a result of them being unimportant causes.

Despite the coherence and superficial plausibility of this position, I am inclined to reject it as a panacea for redundant causation. I do not think it is plausible to suppose that anything which affects the time (or manner) in which an event occurs is automatically a cause of that event. My skepticism remains when I remind myself that we are supposed to be dealing with a "broad and nondiscriminatory" concept of causation. What, after all, would be the point of a concept of causation which counted almost anything prior to an event as a cause of that event? I can see no reason to suppose that such a concept would be important to us, and I conclude that it is not our concept of causation.

In fact, it seems to me that our examples indicate that judgments about what causes what are partly judgments about the importance of the former to the latter. If this is right, trying to analyse a concept of *causation* that is *entirely* free of salience conditions may be compared with trying to analyse a concept of *harm* that is entirely free of salience conditions: it seems plausible to suppose that it is analytically true that there is no such thing as a trivial harm and I suspect that similarly there may be no such thing as a trivial cause.[76]

I will, therefore, accept the Lewisean intuitions that an adequate conception of the fragility of events must meet two criteria:

1) It must allow that events can be postponed (and hastened), and could have happened differently.

2) It must allow for a distinction between causing an event and merely influencing the time or manner in which it occurs.

Together they provide good reason for believing that there are events which could have occurred in ways and at times other than

---

[76] There may of course be more or less important harms and more or less important causes. My suggestion is that there are no unimportant harms or causes.

they do; that is, that we do not have uniformly stringent standards of fragility.

The common suggestion, therefore, seems unsatisfactory. My strategy will differ from it in not denying that there are genuine cases of redundant causation, indeed I will accept that there are. I only deny that there are genuine cases of preemption as defined by Lewis.

Chapter Five

A New Suggestion

Appealing to the fragility of the effect to dissolve asymmetrical redundancy (*i.e.* preemption), without dissolving symmetrical redundancy

Let us suppose then, as I think we should, that Lewis's examples of early and late preemption do, as he says, exhibit genuine redundancy. Let us suppose further, as I think we also should, that the firings of c_1 in these examples do cause the firings of e, while the firings of c_2 do not.

In view of these suppositions, Lewis's position that these are examples of events (namely the firings of c_1) causing distinct events (namely the firings of e), without it being the case that if the causes had not occurred, the effects would not have occurred either, seems unavoidable. Consequently the naive counterfactual analysis seems to be refuted.

In fact, the naive counterfactual analysis can accommodate these suppositions by denying that the description "the firing of e" (and "the death of Edgar") refers univocally. Corresponding to each of

these diagrams there is an event which we may accurately call "the firing of *e*" which is caused by the firing of c_1, and there is another event, just as deserving of the name "the firing of *e*", which is such that, if the firing of c_1 had not occurred it would not have occurred either. There is, however, no event which is both caused by the firing of c_1 *and* is such that it still would have occurred if c_1 had not fired.

Where there is indeterminate reference there is more than one acceptable candidate for being the entity referred to. We may have left the exact boundaries of the spatiotemporal region in which a particular firing of *e* occurs undecided. If so, "the firing of *e*" will be vague and hence indeterminate, but this is not the kind of indeterminacy I have in mind.[77]

It may be that *e* fires on more than one occasion. If so "the firing of *e*" may be ambiguous and hence indeterminate, but this is not the kind of indeterminacy I have in mind either.[78]

---

[77] This kind of indeterminacy is more clearly present in our examples of death. Did I mean brain death? Did I mean heart death? There is no fact of the matter. They are both pretty good candidates.

[78] I anticipate the following objection. If *e* fires more than once, the description "the firing of *e*" is not ambiguous but improper. I reply that there will be contexts in which this description refers to one firing, contexts in which it refers to another, and contexts in

The candidates I have in mind both occur in exactly the same spatiotemporal region.[79] The possiblility of more than one event occurring in exactly the same region is ruled out by a popular view, according to which an event is a region, or the 'content' of a region.[80] Lewis rejects this approach:

> Two events can occur in exactly the same region. An electron's presence in a field can cause its acceleration; radiation of two frequencies can reverberate throughout the same cavity; two chemical reactions can go on in the same flask. More fancifully, there might be goblins made of a sort of matter that passes through our sort without any interaction, and a battle of goblins might occur in the very same region as this conference (the 1981 Conference of the Australasian Association of Philosophy, where this paper was first read).[81]

Those who want to identify events with their regions, or who for whatever reason are committed to denying that more than one event

---

which there is no fact of the matter.

[79] I will follow Lewis in using "region" as an abbreviation for spatiotemporal region.

[80] For example, W.V. Quine *Philosophy of Logic* pp.31-32, and J.J.C Smart "Further Thoughts on the Identity Theory". The distinction between a region and its 'content' does not concern me here. What is important to me about these views, is that they rule out the possibility of more than one event occurring in a single region.

[81] "Events" p. 245

ever occurs in a region are unlikely to change their minds in response to these examples. In some cases they might argue that one of the entities is not a genuine event (an electron's presence in a field?), in others they might argue that the example is too fanciful to be taken seriously (a battle of goblins?), in yet others they might argue that the events do not occur in exactly the same region (two chemical reactions going on in the same flask?).[82] Finally they might argue that some or all of these cases do not really involve two entities at all, but a single entity referred to in two different ways.

Quine adopts the last strategy in defense of the identification of events and regions, against what he takes to be a *prima facie* counterexample, in the following passage:

> Such an account of events does not distinguish between events that happen to take up the same portion of space-time. If a man whistled a song all the while he was walking to the bus stop and not a moment more, then

---

[82] In Lewis's terminology these chemical reactions occur *within* the same region (*i.e.*, the flask), but it is not obvious to me that they occur *in* the same region. Lewis draws the distinction as follows:

> An event occurs within every region that includes the region in which it occurs; and it occurs in the region that is the intersection of all regions within which it occurs. ("Events" p. 244)

If we ignore this distinction, the doctrine that two events cannot occur in the same region will be obviously false.

> presumably the event of his whistling the song and the event of his walking to the bus would both be identified with the same temporal segment of the man. This outcome will be unwelcome if one feels that the event of the whistling and the event of the walking should be distinguished. However, it is not clear to me that they need to be. We still have the general distinction between whistling the song and walking to the bus, because sometimes people do walk to the bus without whistling the song, and even vice versa. Thus we still convey information when we say that the man whistled the song all the say to the bus. We do not trivialize the statement in identifying the events.[83]

The real problem, I think, with identifying these events is that they appear to have different causes and effects. The man's knowledge of the song was presumably a cause of him whistling it, however it was presumably not a cause of his walking. Although the man's arrival at his destination was presumably an effect of his walking, it was presumably not an effect of his whistling.

Quine seems to concede that his proposal is counterintuitive; saying that accepting it requires "imagination".[84] Lewis agrees; claiming that "we do usually think that two different events might

---

[83] W.V. Quine, "Multiplying Entities", in *The Ways of Paradox and Other Essays*, p. 260.

[84] ibid.

occur in the very same region".[85] I do not doubt that 'we' would ordinarily allow for the logical possibility of event cohabitation. I do not think, however, that there are any intuitively compelling 'real world' examples of this phenomenon.[86] After all, it is not at all clear that the whistling and the walking, in Quine's example, do occur in exactly the same region. Arguably they could be identified with different parts of the same temporal segment of the man: arguably the whistling does not occur in the region of his legs, and the walking does not occur in the region of his mouth.

I will argue, however, that Lewis's own theory is rightly committed to the position that the phenomenon of more than one event occurring in exactly the same region is not only possible, it is commonplace. In particular, I will argue that a plurality of events occur in the region in which neuron *e* fires (and the region in which Edgar dies) in our preemption examples. These events provide us with the candidates which create the indeterminacy of which I

---

[85] "Events" p. 246

[86] Lewis may have felt the same way. This is suggested by the fact that he illustrates this claim about what we usually think by appeal to his 'battle-of-goblins' example, rather than any of the more realistic examples which were supposed to illustrate the same point.

speak. I will not argue, however, that this is what we usually think. My position is, to a certain extent, counterintuitive. I will argue, however, that the intuition with which it is in conflict is untenable, and can be explained away.

Postulating an indeterminacy would, of course, be outrageous, if it were motivated only by a desire to deal with what would otherwise be counterexamples to my analysis. I will argue, however, that we should be committed, independently of preemption problems, to the position that expressions like "the firing of *e*" are indeterminate in just the required way. I will call this kind of indeterminacy *modal indeterminacy*.

The Profligate Theory of Events:

Lewis points out that in each of the examples which he presents as illustrating the possibility of two events occurring in exactly the same region "it would have been possible for one of the two events to occur without the other."[87] Presumably Lewis would be prepared to make a stronger claim about these examples: not only

---

[87] "Events" p. 245

would it have been possible for one of the two events to occur without the other, it would have been possible for *either* of the two events to occur without the other.[88] The claim that an event could have occurred without its 'regional partner' is susceptible to two interpretations. Firstly, it could mean that there is a possible world in which it occurs, but its partner does not. Secondly, it could mean that there is a region of a possible world in which it occurs, but its partner does not. Lewis makes it clear that he has the latter interpretation in mind in the following passage:

> ... I can think of no plausible case of two events such that, necessarily, for any region, one occurs in that region iff the other does. Two such inseparables would be causally indistinguishable on a counterfactual analysis of causation, so it is hard to see how my treatment of causation could need them both. I shall therefore take it that for any two events there is some region of some world where one occurs and the other does not.[89]

Although Lewis dismisses the possibility of two events being *inseparable*, in this sense, from each other, he does not rule out the possibility of one event being *inseparable* from another: in Lewis's

---

[88] We could say the same thing of Quine's example.

[89] ibid

terminology the former event would *imply* the latter event in the following sense:

> Let us say that event *e* *implies* event *f* iff, necessarily, if *e* occurs in a region then also *f* occurs in that region.[90]

In other words, in the language of possible worlds, any world in which *e* occurs is a world in which *f* occurs in the same region as *e*. This definition requires, not only that *e* be inseparable from *f* in the sense that *f* occurs in each possible world in which *e* occurs, but also in the sense that *f* occurs in each of the possible regions in which *e* occurs.

Lewis tells us that he has changed his mind about whether there are any instances of this relation. His earlier view was that all events are separable from each other, *i.e.*, that no event implies any other event. I will call Lewis's earlier view, *the identity thesis*, since it entails that *e* implies *f* iff *e* is *f*.[91] He came to think, however, that some events are inseparable from other events, *i.e.*, that some events imply other events. Following Michael McDermott's

---

[90] "Events" p. 255

[91] I borrow the terminology from Alvin Goldman. Although it is not explictly defined by him, it is consistent with his usage, see Alvin I. Goldman (1970) p. 3.

terminology, I will call the denial of the identity thesis "the profligate theory of events".[92] We may say that Lewis's considered view is that there are two kinds of cases of pairs of events occurring in the same region: those in which the events are mutually separable, and those in which one of the events is inseparable from the other.

Lewis does not try to offer any intuitively compelling examples of the latter kind of pair: nor does he suggest that we usually think that there are any. In fact, he originally rejected the existence of such pairs, because putative examples of them seem not to be examples of two events, but of one. Lewis came to change his mind, not because of his views about events (at least not directly), but because of his views about causation.

Lewis's later view is introduced by means of an example from Alvin Goldman:

---

[92] "Redundant Causation" p.p. 531-532. McDermott does not explicitly define this expression, but I think my definition conforms closely to his usage. McDermott does not like the profligate theory. He claims that "Lewis seems to find the profligate theory of events an embarrassment". I don't believe this is accurate. Although I do think that Lewis's discussion of preemption fails to do justice to its implications.

> John answers the phone and says "hello." He says "hello" because he wishes to greet the caller. But John has just been quarreling with his wife and is in a tense emotional state. As a result, he says "hello" very loudly. He doesn't intend to shout over the phone; it just comes out that way.[93]

Lewis suggests the possibility that (at least) two events occur in the region in which John says 'Hello':

> Arguably there is one event that occurs which is essentially a saying-"Hello" and only accidentally loud; it would have occurred even if John had spoken softly. Arguably there is a second event that implies, but is not implied by, the first. This event is essentially a saying-"Hello"-loudly, and it would not have occurred if John had said "Hello" but said it softly.[94]

Lewis came to think that it was not just arguable that both these events occur: it is true. If they do both occur, the second event has a richer essence than the first. For this reason I will call the first alleged event "Robust", and the second alleged event "Fragile". In general we may say - in the language Lewis uses in his discussion of redundant causation - that if one event implies another, the implying event is more fragile than the implied event.

---

[93] Alvin I. Goldman, (1970) p. 3

[94] "Events" p. 255

Lewis argues that "we need both events" because "they differ causally".[95] It is certainly true that according to counterfactual analyses of causation (both my own and Lewis's) Robust and Fragile would differ causally if both occur. If not for John's tension, he would have spoken in his normal soft voice. Consequently if Robust occurs in the actual world, it would still have occurred under this counterfactual supposition, whereas Fragile would not have ocurred (whether or not it actually does). So *if* both Robust and Fragile actually occur, the counterfactual analysis will count the latter, but not the former, as an effect of John's tension. This is not yet, however, a reason for believing that they do both occur. Lewis says "an adequate causal account of what happens cannot limit itself to either one of the two."[96] But why should we accept this condition of adequacy?

Lewis's argument for the profligate theory, if I understand it correctly, can be seen, at least in part, as an extension of the method employed in his too-many-causes argument against 'the common suggestion'. This involved inferring the essence of an event

---

[95] ibid.

[96] ibid.

from our beliefs about what caused it. I have accepted Lewis's position that we do not want to say that anything which has only a slight influence on the time or manner of occurrence of an event is thereby a cause of that event. I have also pointed out, however, that with each of the examples used to illustrate this point there are conceivable contexts in which we would be inclined to count such noncauses as genuine causes after all.

We could accommodate this variability by appealing to a variability in what we mean by 'cause'. For example, we could take it to be the result of a distinction between a concept of causation with salience conditions and a concept of causation without salience conditions. I have already argued against this position.

A better approach to this variability is to attribute it to the relata, rather than the relation. This approach is suggested by Lewis's development of the example mentioned earlier, in which we were asked to agree that Boddie's eating does not consititute a cause his death, except in "a special context". He gives the following example of such a context:

> Pleased that Boddie is dead but horrified that the death was lingering, the poisoner says: if only he hadn't eaten, *this* wouldn't have happened - and by "this" he means the

> death, taken as very fragile. Maybe indeed that context makes it right to say that the eating caused the death.[97]

Menzies seems to be making this point when he says that sometimes we want to "state the causes of a fragile event, say, a person's slow, lingering death, but we can also state the causes of his death *simpliciter*."[98]

If we can postulate the occurrence of both an event 'taken as very fragile', and that same event *simpliciter*, we should also, I submit, be free to postulate the occurrence of an event 'taken as very robust': "It was bound to happen some day", says the fatalist, expressing his indifference to Boddie's inevitable, if premature, death. Boddie's death, taken as very fragile, implies Boddie's death *simpiciter*, and both these events imply his death, taken as very robust.

We can now see that the method of inferring the fragility of events from our judgements about their causes leads, not only to the rejection of 'extreme fragility', but also to the need for events which imply other events. Any adequate theory of causation must

---

[97] "Postcript E", p. 198

[98] Peter Menzies, (1989) p. 650

have a place for our capacity to express causal propositions of differing degrees of specificity. The profligate theory of events meets this need, for counterfactual analyses.

Jonathan Bennett has expressed concern over the way in which Lewis refers to events by describing their essences, rather than by using event nominals drawn from natural language; complaining of "the estrangement between Lewis's events and ordinary event names".[99] He expresses uncertainty about what causal relations a pair of events like Robust and Fragile would stand in:

> ... I have never before met these two putative events, or anything like them, and I don't know how to behave in their presence. I recognize them as products of Lewis's theory, but not as parts of my everyday conceptual armory.[100]

Lewis makes no secret of the fact that he is not concerning himself with the relation between events and the means by which we ordinarily refer to them:

---

[99] Jonathan Bennett (1988). Lewis calls Robust "the weak event" and Fragile "the strong event", Bennett calls Robust "General" and Fragile "Specific".

[100] ibid.

> "When I introduce nominalisations to denote events, as I shall, it will not be analysis of natural language but mere stipulative definition."[101]

Unlike familiar events, such as deaths, neuron firings, and people saying 'Hello', we seem to have no intuitive grasp of the causes and effects of Fragile and Robust. I have some sympathy then for Bennett's complaint that Lewis does not "connect his theory with everyday thought and talk."[102] But this does not mean that it cannot be done. I will try to go some way towards establishing some such connections.[103]

One possible response to Bennett's concern would be to adopt Goldman's position that Robust is familiar to us as the referent of the everyday event nominal "John's saying 'Hello'", and Fragile is equally familiar as the referent of "John's saying 'Hello' loudly".[104]

---

[101] "Events" p. 241

[102] Jonathan Bennett (1988) p. 68

[103] In what follows I am particularly heavily indebted to discussions with David Lewis.

[104] What is not so familiar, of course, is the idea that John's saying "hello", and John's saying "hello" loudly, are different events.

That is, 'in the material mode', Robust is John's saying 'Hello', and Fragile is John's saying 'Hello' loudly.

These identifications will be tempting if, like Goldman, one feels that John's saying 'Hello' is not caused by his tension, whereas John's saying 'Hello' loudly is caused by it. There is surely something to Goldman's position. It illustrates how our intuitions about the causes of what we would be inclined pretheoretically to call a single event seem to vary, depending on how we refer to it, and in particular how much detail is included in the event nominal.

Lewis does not, however, assume that the more detailed description must denote the more fragile event, or that the less detailed description must denote the more robust event:

> Indeed, even if there are two different events, it still does not follow that one description denotes one and the other denotes the other. If both descriptions are somewhat vague or ambiguous, it could be that both denote both.[105]

Bennett has pointed out that Lewis's position has the following consequence:

> Now, consider someone who asserts "John's saying 'Hello' caused Fred's reply". He could permissibly mean that Fred wouldn't have replied if John hadn't said "Hello" so loudly

---

[105] "Events", p. 255

> or that Fred wouldn't have spoken if John hadn't said "Hello" in some fashion, and it might be that in one meaning the remark is flatly false and in the other perfectly true.[106]

Lewis could, and presumably would, respond that this assertion may be ambiguous in this way, and if so it is a virtue of his theory that it agrees.[107] In fact I think this assertion is ambiguous in this way, as does everyone I have discussed the example with.

If we accept this ambiguity we should also accept that someone who asserts "John's saying 'Hello' was caused by his tension" could permissibly mean that his tension caused him to say 'hello' so loudly, or that it caused him to say 'Hello' in some fashion: in the former meaning the remark is true, whereas in the latter meaning it is false. Again, it seems to me that this assertion can be ambiguous in this way; again, everyone with whom I have discussed this example agrees. So it appears that we should not accept, without qualification, Goldman's claim that John's saying 'Hello' is not caused by his tension; there seems to be a perfectly good sense in which it is.

---

[106] Bennett (1988) p. 68

[107] Lewis has recently explicitly confirmed that he does think this assertion may be ambiguous in this way.

Causal statements which use imperfect nominals like "John's saying 'Hello'" (or "the firing of neuron *e*") tend to be rather awkward, and often seem to cry out for clarification. Although my first reaction is to agree with Goldman that John's tension was not a cause of his saying 'Hello', in the circumstances he outlines; a context in which the speaker was particularly interested in the loudness of what John said would incline me somewhat toward an interpretation according to which it may be true. A speaker could make it clear that she has such an interpretation in mind by emphasizing the word 'Hello'; pronouncing it noticeably louder than the rest of the words in the assertion.[108] I think that another way of clarifying this statement would be to replace the ungainly imperfect nominal with a noun-infinitive. We could then distinguish two ways of understanding it:

1) John's tension caused him to say 'Hello'.

2) John's tension caused him to say 'Hello' too loudly.

The first statement seems false, the second seems true.

---

[108] This was pointed out to me by Lewis.

Bennett agrees that noun-infinitives have this characteristic.[109] However, he takes this to be because noun-infinitives do not refer to events at all, but to facts. Bennett believes that there are two kinds of causation 'fact causation' and 'event causation', and that the former is superior precisely because it allows us to express differing kinds and degrees of specificity.[110] Bennett does not think that this distinguishes the noun-infinitive form from the imperfect nominal form, however, since he believes that imperfect nominals also refer to facts rather than events.[111]

If, as Bennett claims, 'event causation' statements are inferior in this way, one would expect to see this reflected in our intuitive judgments about their truth values. But, I do not think they are. Nor does Bennett suggest that they are. Rather, he argues for this view of event causation by arguing that the conception of events which it

---

[109] Bennett, (1988) p. 25

[110] ibid. p. 139. For another advocate of this view see D. H. Mellor, (1995) pp. 156-192.

[111] Bennett, (1988) p. 7

requires (*i.e.*, the profligate theory) is mistaken; that it is committed to a kind of double-counting.[112]

Double-counting?

The profligate theory has seemed to many philosophers, including Lewis for a while, to involve a multiplication of entities, if not beyond necessity, then beyond commonsense. It appears, for example, to be committed to more greetings, and more deaths of Boddie, in the examples I have just been discussing than pretheoretically we thought there were.

Bennett discusses an example in which Peter leaves a party because he dislikes the music. His departure is noisy because of an incidental coughing fit:[113]

---

[112] It is difficult to understand why he did not use this argument as a direct criticism of Lewis's commitment to the existence of both Fragile and Robust.

[113] Bennett's conviction that events cannot be individuated finely leads him to a strikingly counterintuitive view about this example. He asks us to consider the following two statements:
1e. The music caused Peter's departure.
2e. The music caused Peter's noisy departure.
Pretheoretically, I think, we want to say that 1e is true and 2e is false. Bennett claims, however, that they are both true.

> On any viable theory of events, a person who leaves only once is the subject of only one departure, and so "his departure" and "his noisy departure" refer to the same event, with the latter describing it more fully ...[114]

Athough the profligate theory is not committed to denying that, at least in some contexts, "his departure" and "his noisy departure" can refer to the same event, it does seem committed to the *prima facie* strange view that Peter departed more than once, *i.e.*, that he was the subject of more than one departure.

According to the profligate theory of events there may be "so to speak, a more and a less detailed version of what happens" in the region of Peter's departure, both of which are events.[115] Which of them is the one departure that pretheoretically we thought we had? Any answer seems arbitrary. Accepting that there is more than one departure seems absurd.

---

[114] ibid. p. 25. There is a similar argument, using a different example, on pp. 10-11.

[115] Lewis refers to a pair of events related by implication in this way in "Events", p. 257.

This is, or is at least closely analogous to, Peter Unger's 'problem of the many'.[116] Lewis discusses this problem with the help of an example from Peter Geach:[117]

> Cat Tibbles is alone on the mat. Tibbles has hairs h_1,h_2, ..., h_{1000}. Let c be Tibbles including all these hairs; let c_1 be all of Tibbles except for h_1; and similarly for c_2, ... , c_{1000}.[118]

After rejecting, for reasons that need not concern us here, Geach's argument that this information alone generates a paradox, Lewis continues the story:

> Tibbles is shedding. When a cat sheds, the hairs do not come popping off; they become gradually looser, until finally they are held in place only by the hairs around them. By the end of this gradual process, the loose hairs are no longer parts of the cat. Sometime before the end, they are questionable parts: not definitely still parts of the cat, not definitely not. Suppose each of h_1,h_2, ..., h_{1000} is at this questionable stage. Now indeed all of c_1, c_2, ... , c_{1000}, and also c which includes all the questionable

---

[116] Peter Unger, (1980)

[117] Peter Geach, (1980) pp. 215-216.

[118] David Lewis, "Many, but Almost One" in *Ontology, Causality and Mind: Essays in Honour of David Armstrong* (Cambridge, 1993) edited by John Bacon, Keith Campbell, and Lloyd Rheinhardt, p. 24. Henceforth, "Many, but Almost One".

> hairs, have an equal claim to be a cat, and equal claim to be Tibbles. So now we have a thousand and one cats.[119]

Which, of course, is a thousand more than we thought we had. Which is the one cat we thought we had? Any answer seems arbitrary. Accepting that there is more than one cat seems absurd. I follow Lewis in thinking there are two acceptable solutions to this paradox. One is to accept its 'absurd' conclusion, *i.e.*, that there really are 1001 cats on the mat, and find a way to make this seem less counterintuitive. I will discuss this first.

Lewis argues that the real opposite of identity is not difference (*i.e.*, non-identity), but distinctness (*i.e.*, non-overlap).[120] We have a range of cases: at one end we have identity, at the other end we have distinctness, and in between we have various degrees of partial identity. Siamese twins, sharing one finger, are almost distinct.[121] Any two of the candidates for being the cat on the mat are almost identical:

---

[119] ibid. p. 25

[120] He attributes the idea to David Armstrong, see *Universals and Scientific Realism*, (Cambridge, 1978) Vol. 2, pp. 37-38.

[121] The example is Lewis's, "Many, but Almost One", p. 33.

> Remember how we translate statements of number into the language of identity and quantification. 'There is only one cat on the mat' becomes 'For some x, x is a cat on the mat, and every cat on the mat is identical to x'. That's false, if we take 'identical' to express the complete and strict identity that lies at the end of the spectrum. But the very extensive overlap of the cats does approximate to complete identity. So what's true is that for some x, x is a cat on the mat and every cat on the mat is almost identical to x. In this way, the statement that there is one cat on the mat is almost true. The cats are many, but almost one. By a blameless approximation, we may say simply that there is one cat on the mat. Is that true? - Sometimes we'll insist on stricter standards, sometimes we'll be ambivalent, but for most contexts it's true enough.[122]

One difference between the plurality of cats in this example, and the plurality of departures (or deaths, or greetings) postulated by the profligate theory of events is that the former obviously share common parts, whereas it is far from obvious whether there is any sense in which the latter do. If partial identity requires an identity of parts, it appears not to help the profligate theory.[123]

---

[122] ibid. pp. 33-34

[123] Although Lewis does seem to think this, Armstrong believes that states of affairs, at least, can be partially identical without sharing a common part. See his reply to "Many, but Almost One" p. 40.

In fact, Lewis's theory of events is *prima facie* well suited to the profligate theory of events; because according to it, if one event implies another, the former is literally a part of the latter.

We have seen that Lewis does not identify events with regions. For Lewis, an event is a class of possible regions, no more than one member of which is in any possible world; it occurs if and where and when there is a region that is a member of it.[124] If this is right, then Boddie's death, taken as very fragile, and that death, taken as very robust, for example, though not identical, are not distinct either. The former is a subclass of the latter, and therefore a part of it.[125] He has the following reponse to the intuition that the profligate theory is committed to double-counting:

> Compare the equally persuasive intuition that it is double-counting to include both atoms and molecules in our inventory of being - an adequate answer is that the molecules and their atoms are not distinct.[126]

---

[124] "Events", p. 245. Lewis does not think that every such class of possible regions is an event.

[125] For Lewis's defense of the view that subclasses of a class are parts of that class see his *Parts of Classes*, (Oxford, 1991) pp. 3-6.

[126] "Events" p. 256

It is not immediately clear, however, even if we were to accept that the fragile 'version' of Boddie's death, is literally a part of the 'robust' version of Boddie's death, why that should allay the suspicions of those who think that by being committed to both of them we "count something twice over."[127] Presumably no one would be tempted to identify atoms and molecules; they really are two things, and what's more they don't even appear to be almost identical.

So even if we were to accept the ontological commitments of Lewis's theory of events, *i.e.*, the existence of merely possible and hence unactual regions, it is not clear that they provide him with an adequate response to the accusation of double-counting.

To see why not consider another example from Lewis's discussion of the 'problem of the many':

> Fred's house taken as including the garage, and taken as not including the garage, have equal claim to be his house. The claim had better be good enough, else he has no house. So Fred has two houses. No![128]

---

[127] ibid.

[128] "Many, but Almost One", p. 29

In this case though, Lewis argues that the almost-identity approach does not work:

> But although the two house-candidates overlap very substantially, having all but the garage in common, they do not overlap nearly as extensively as the cats do. Though they are closer to the identity end of the spectrum than the distinctness end, we cannot really say theyr'e almost identical. So likewise we cannot say that *the two houses are almost one.*[129]

Similarly, even if we agreed that Boddie's death, taken as very fragile, overlaps extensively enough with Boddie's death *simpliciter* to make them almost identical, we should feel some reluctance to say that the former is almost identical with his death, taken as very robust.

Lewis argues that there is another solution to the 'problem of the many', which works better than the almost-identity approach for 'Fred's two houses':

> You say that a famous architect designed Fred's house; it never crossed your mind to think whether by 'house' you meant something that did or that didn't include the attached garage; neither does some established convention or secret fact decide the issue; no matter, you knew that what you said was true either way.[130]

---

[129] ibid. pp. 35-36

[130] ibid. p. 29

Following van Fraasen's strategy for dealing with unmade semantic decisions, we may say a statement is *super-true* iff it is true on all interpretations; it is *super-false* iff it is false on all interpretations; if it is true on some interpretations and false on others it has a super-truth-value-gap.[131]

Super-truth occupies the role in a less than fully interpreted language that truth *simpliciter* occupies in a fully interpreted language: it is that at which honest speakers aim. So we are entitled to say that Fred has only one house, because it is super-true that Fred has only one house, *i.e.*, it is true on each interpretation of 'house'. Similarly, the profligate theory can agree with commonsense that Boddie is the subject of only one death, since it is super-true, ie. true on each interpretation of 'Boddie's death', that there is only one of them.

There is often a significant lack of constancy in our judgments about the causes of a given event. Some, but not all, of this inconstancy can be attributed to differing resolutions of an indeterminacy about the region in which the event occurs. For example our judgments about the causes of Boddie's death might

---

[131] Bas C. van Fraassen, (1966)

vary, depending on whether we are thinking of heart death or brain death.

There is a residual inconstancy, however, which has inclined many philosophers to deny that causation is a relation between events. It might instead be a relation between aspects of events.[132] Or it might be a relation between events-in-contrast-to-alternatives.[133] Or it might be a relation between facts.

I think that all these moves are unnecessary. We can instead say that what is in one sense a single event, is another sense many - one for each of the judgments about 'its' causes that we want to make (and maybe some that we do not want to make on pragmatic grounds). This inconstancy is due to semantic indecision, not about the region in which the event occurs, but about the essential properties of what occurs in a particular region. If I am right the 'paradox' created by the profligate theory of events is just a new example of the old 'problem of the many'. The intuition that there is only one event can be respected by arguing that it is almost true, or

---

[132] See L.A. Paul, "Aspect Causation", forthcoming

[133] Christopher Hitchcock, "The role of Contrast in Causal and Explanatory Claims", *Synthese* (107, 1996) pp. 395-419

that it is super-true.[134] Either way, I have seen no reason to believe that the profligate theory of events gives us any problems we did not already have: a new example of an old problem is not a new problem.

In support of the profligate theory of events and the existence of modally indeterminate descriptions, it is worth pointing out that closely related phenomena are already to be found in probability theory, in which an event is identified with a set of possible outcomes of a probability experiment.[135] To speak of 'the outcome' of a particular experiment may well be to speak ambiguously. Several events, some implying others, may be referred to in this way. If, for example, the experiment is the throwing of a die, the speaker may be referring to the relatively fragile outcome of getting a two or the relatively robust outcome of getting an even number, even though pretheoretically we would want to say this is a single event. If the profligate theory of events is committed to an invidious double-counting, so it seems is standard probability theory.

---

[134] Or these solutions can be combined, see Lewis, (1993) p. 35

[135] Although Lewis does draw a distinction between sets and classes, this distinction makes no difference to his theory of events.

A philosophical precedent for the profligate theory can be found in J. L. Austin's *How to Do Things with Words*, in which he distinguishes between a variety of acts (*e.g.*, locutionary, illocutionary, and perlocutionary) all of which would be, according to the identity thesis, a single act.

A friend of the identity thesis could reply that this only shows that the word 'event' is misused in probability theory, and by J. L. Austin.[136] Lewis has said that although there are many different conceptions of events, his "interest is in events as causes and effects".[137] Being a Lewisean event is neither necessary nor sufficient for being an entity suitably called an 'event' in ordinary language. It is not sufficient because there are entities which we may call 'events' which are not, and could not be, causes or effects of anything:

> A certain mathematical sequence converges. There is some entity or other that we may call the converging of the sequence. The sequence converges rapidly iff, in some sense, this entity is rapid. I have no objection to that; but I insist that the converging of the sequence,

---

[136] Austin and I both assume that an act is a kind of event.

[137] "Events", p. 241

> whatever it may be is nothing like any entity that causes or is caused.[138]

It is not necessary, because there are entities which it is wrong, in ordinary language, to call 'events', but which do have causes and effects. For example, in ordinary language the term 'event' implies change, but Lewis does not restrict the concept in this way:

> Terminology is not the issue. If it is abuse of language to call unchanges "events," so be it. The point is that we must have them as causes and effects.[139]

Lewis claims that events "are not much of a topic in their own right", their importance is derived from their role "in the discussion of other topics."[140] One such topic is causation, another is probability: both require a 'profligate' conception of events.

We have seen how one of Lewis's arguments against extreme fragility can be turned into an argument for the profligate theory of events. His other argument against extreme fragility seems to support the same conclusion. We do often make counterfactual suppositions according to which an actually occurring event occurs

---

[138] ibid.

[139] ibid. p. 261

[140] ibid. p. 241

at a somewhat different time, or in a somewhat different manner. However we can also often, perhaps always, make a qualitatively identical supposition according to which it is a different event which occurs at that time and in that manner. Lewis gives the following example of this phenomenon:

> You can say: the performance should have been postponed until the singer was over his laryngitis; then *it* would have been better. You can just as well say, and mean nothing different: the performance should have been cancelled, and another, which would have been better, scheduled to replace it. There's no right answer to the question how fragile the performance is.[141]

Or rather, there is more than one right answer, because there is a fragile performance, which could not have happened very differently, and there is a robust performance which could. Both are permissible referents of "the performance".

The profligate theory of events should make us reconsider our preemption problems. This is so because it opens up the possibility that the descriptions we have been using to denote the effect in these examples could be modally indeterminate in the way required to dissolve the problem.

---

[141] "Postscript E", p. 197

Chapter Six

Late Preemption Reconsidered

Consider an analogue of Lewis's reasoning about the causes of John's saying hello, applied to the firing of neuron *e* in Diagram 3. When this neuron fires at time *t* at least two logically related events occur. One of them is essentially a firing of *e* and only accidentally a firing of *e* at *t*. The other is essentially a firing of *e* at *t*. The real reason why we need both events is that they differ causally. The former is not caused by the firing of c_1, since if c_1 had failed to fire it would have happened anyway. The latter is caused by the firing of c_1, since if c_1 had failed to fire it would not have happened. So we seem to have just as good a reason to distinguish between different candidates for being *e*'s firing as we do between different candidates for being John's 'Hello'.

One difference between the candidates for being John's 'Hello' and the candidates for being the firing of *e* in these examples is that the former differ in their *manner of occurrence* fragility whereas the latter differ in their *temporal* fragility. But I do not see how

this difference could be relevant. Just as an adequate causal theory should be capable of fine-graining causal propositions with respect to the manner of occurrence of the effect, it should also be capable of fine-graining causal propositions with respect to the time of the effect. We have already discussed such temporal fine-graining in connection with Boddie's death.

How does this leave our initial intuitions about the example? It has seemed obvious to every one I know of who has written about this subject that the firing of c_1 definitely is a cause of the firing of *e*. I don't think this is obvious at all. I am not, of course, denying that the former firing caused the latter. Rather I think there is a sense in which it did, and a sense in which it did not.

The question "Did the firing of c_1 cause the firing of *e*?" should, I submit, be met with the response "Are you asking whether the firing of c_1 caused *e* to fire at the time it did, are you asking whether it caused *e* to fire as early as it did, or are you asking whether it caused *e* to fire at all?"[142] We should answer the first two questions in the affirmative and the third in the negative. If I

---

[142] There may well be other ways of understanding the question. But these three are particularly relevant to our discussion.

am right late preemption is not preemption at all; depending on how fragile you take the effect to be, it is either staightforward causation or symmetrical overdetermination.

It may seem that this approach does not do justice to the widespread intuition that the firing of c_1 is a cause of the firing of *e*. I think it does, because not only does it grant that there is a sense in which this intuition is accurate, it is also consistent with the position that anyone who asserts that the firing of c_1 is a cause of the firing of *e* is asserting something true. In defence of this claim I don't think I can do better than quote Lewis on a related subject:

> There is a rule of accommodation: what you say makes itself true, if at all possible, by creating a context that selects the relevant features so as to make it true. Say that France is hexagonal, and you thereby set the standards of precision low, and you speak the truth; say that France is not hexagonal ... and you set the standards high, and again you speak the truth. In parallel fashion, I suggest that those philosphers who preach that origins are essential are absolutely right - in the context of their own preaching. They make themselves right: their preaching constitutes a context in which *de re* modality is governed by a way of representing ... that requires match of origins.[143]

---

[143] *On the Plurality of Worlds*, pp. 251-252

In parallel fashion, I suggest that the philosophers who say that the firing of c_1 causes the firing of e are right - in the context created by what they say. Thus I think I do justice to this widespread intuition, despite there being a sense in which it is mistaken.

We also have the intuition that the firing of c_2 is not a cause of the firing of e. I once thought that I could easily accommodate this intuition by adopting the position that no matter how fragile you take the firing of e to be it is not caused by c_2. After all we have seen no reason to believe that e would have occurred at a different time or in a different manner, if c_2 had not fired.

Conversations with Jonathan Schaffer have led me to reconsider this assumption. It may be that c_2 has some slight affect on the time or manner of e, and if so I do not think it would radically change our intutions about the example. So it seems that we should after all accept that there may be a sense in which c_2's firing is a cause of e's firing.

I do not think this is a problem. There is no danger of losing the asymmetry which all agree is a feature of these examples. The firing of c_1 and of c_2 may both be causes of the firing of e under

extremely stringent standards of fragility, neither will be causes under extremely robust standards of fragility, but under ordinary standards of fragility the firing of c_1 will be a cause and the firing of c_2 will not. I think this does justice to the widespread intuition that c_1's firing is, and c_2's firing is not, a cause of e's firing *simpliciter*.

Lewis writes that there are two specific problems with fragility solutions to late preemption problems. The first is that "we need some definite rule to tell us when we should raise the standard: when is dependence among fragile versions relevant, and when is it not?"[144] Well we have the definite rule that the firing of c_1 is a cause of any version of the firing of e fragile enough not to be able to occur after the time it did occur, and I've seen no reason to think that this does violence to any legitimate causal intuitions.

We don't need to be temporal essentialists (one way or otherwise) to believe that there are candidates for being the firing of e which are this fragile. It might be objected, however, that although I have avoided the implausible postion that *every* event is this fragile, there is a residual implausibility in supposing *any* event

---

[144] "Postcript E", pp. 204-205

could be *this* fragile. Lewis has appealed to an intuition that there are limits to the fragility of events:

> There is an event which is accidentally a fiddling while Rome burns, but I doubt that any event is essentially a fiddling while Rome burns. And the example can be made even more extreme. There is an event which is accidentally classifiable as follows: it is a fiddling in the presence of a boy whose grandson will first set foot on the moon. Surely no event is essentially *that*![145]

Whatever we may think of the intuition Lewis is appealing to here, once we accept the profligate theory of events, I can see no reason to place any restrictions on how temporally fragile events can be.[146] To do so would be to place unnecessary limits on how temporally fine-grained causal propositions can be.

We surely want our theory of causation to permit us to ask of any event "What made it occur *at the time it did*?" In order to do so we will need a *candidate* for being that event which essentially occurs at the time it does. Consequently I embrace the principle that every event is implied by an event which essentially occurs at the

---

[145] "Events", p. 258

[146] I will defer discussion of how *manner of occurrence* fragile events can be, until I discuss early preemption. This issue will be relevant to my discussion of it in a way that is closely analogous to the way in which the issue of how *temporally* fragile events can be is to my discussion late preemption.

time it does.[147] For anyone who has a metaphysical intuition that no events are this fragile, or not enough to make this principle true, we could replace it with the weaker principle that every event is implied by an event which essentially occurs no later than it actually does. Although the weaker principle would be sufficient to deal with all cases of late preemption, since in these cases the counterfactual 'effect' must take place after the actual one, I see no reason not to advocate the stronger and simpler version of the principle.

Lewis says the second problem with fragility approaches to late preemption is that they may give the wrong answer:

> Let c_1 be a preempting cause of e, and let c_2 be the preempted alternative, in a case of late preemption. Without c_1, e would have been delayed; and so a more fragile version of e would not have occurred at all. So far, so good. But it may also be that some side effect of c_2 substantially influences the time and manner of e; in which case, unfortunately, a version of e that is fragile enough to depend on c_1 may depend on c_2 as well. Indeed, it may take more fragility to give us the dependence on c_1 that yields the right answer than it does to give us the dependence on c_2 that yields the wrong answer.[148]

---

[147] One way of being implied by such an event of course is to be identical with it.

[148] Lewis, "Causation", p. 205

Michael McDermott illustrates this argument with his own example:

> Two shots are fired in quick succession. The first bullet perforates Jones, the second passes through the hole in its wake, scarcely touching the sides ... Suppose that the second shot makes a particularly loud noise; this causes Jones to twitch slightly (while the first bullet is still on its way), so that the first bullet (with the second in its wake) perforates a somewhat different part of his body; this makes him die more (or less) quickly. Without the first shot, he would have died at a *slightly* different time; without the second, he would have died at a *substantially* different time; if the first shot was a cause of his death because of the death's rich essence, so was the second. But common sense is still clear that the first shot was the sole cause of Jones's death.[149]

But I don't think common sense is clear about this at all. If the sound of the second shot causes Jones to move in such a way that he dies substantially quicker then he otherwise would, surely it has at least as much right to be called a cause of Jones's death as the first shot does. The second shot should count as a cause of Jones's death even if the second bullet does not. If the sound of the second shot causes Jones to move in such a way that he dies substantially slower, we would be less likely to count it as a cause; which is what we would expect given the above discussion about delayed deaths.

---

[149] Michael McDermott, "Redundant Causation", p. 530

Chapter Seven

Early Preemption Reconsidered

We cannot use exactly the same strategy for Diagrams 1 and 2 that we used for Diagram 3, because there need not be any time difference between the actual firing of *e* and the one which would have occurred had c_1 not fired. But it may well be that, in our world at least, there will always in fact be some difference between them which would open the door to dissolving putative cases of early preemption by a similar method to that which I applied to putative cases of late preemption.

In order to be sure that such a move is as widely applicable as possible, we will need a candidate for being the firing of *e* which essentially occurs at the time and in the *manner* it does. In view of this I embrace the principle that every event is implied by an event which essentially occurs at the time and in the manner it does.

Will this principle be sufficient to deal with all cases of early preemption? It seems fairly clear that it will be sufficient to deal with most examples of this phenomenon discussed in the literature,

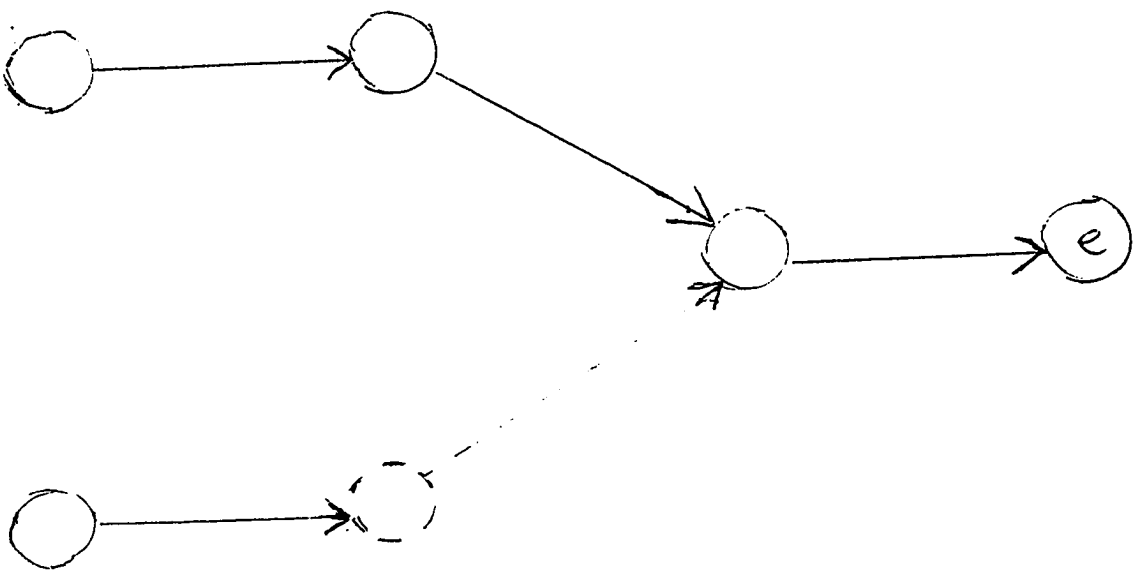
including my own. In the neuron diagrams that Menzies, Ramachandran, Hall, and myself use to illustrate early preemption neuron *e* is what I have called 'the merging neuron'. It is plausible to suppose that in such cases there will always be some difference in manner between the actual and counterfactual firings of *e*, since *e* would have been stimulated from a different direction by the unactualized process; consequently energy would have flowed through it in a different direction. Furthermore, in my other example of early preemption involving a preempted killing, it is hard to imagine how the actual and counterfactual killings would have been exactly the same in manner: they would at least have involved different bullets.

The problem with this simple dissolution of the problem is that neuron *e* need not be the merging neuron. Neuron *e* can be, as it is in Lewis's original diagrams, a neuron which is caused to fire by the firing of the merging neuron. An example of this is illustrated by the following extension of Diagram 1:[150]

---

[150] The difference between diagrams 1 and 2, *i.e.*, the presence in the former and absence in the latter of an intermediate neuron between the branching and merging neuron, is irrelevant to my account.

Diagram 5:



The same difficulty could be created for my macroscopic example of early preemption by changing the issue from the causes of Edgar's death to the causes of his funeral. In cases of this kind it seems harder to justify the claim that the manner of the actual and counterfactual firings of *e* need to be even slightly different. Now we really do seem to have a problem for the naive counterfactual analysis of event causation.

Or at least we have such a problem, if we accept the Lewisean criterion for the trans-world identity of events, *i.e.*, Lewis's claim that "difference in time or manner is what it takes to make a numerical difference between an event that actually occurs and one

that would have occurred under some counterfactual supposition".[151]

This suggests partisanship in a debate which David Kaplan has identified in the following famous passage:

> There seems to be some disagreement as to whether we can meaningfully ask whether a possible individual that exists in one possible world also exists in another without taking into account the attributes and behavior of the individuals that exist in the one world and making a comparison with the attributes and behavior of the individuals that exist in the other world. The doctrine that it does make sense to ask - without reference to common attributes and behavior - whether *this* is the same individual in another possible world, that individuals can be extended in logical space (*i.e.* through possible worlds) in much the way we commonly regard them as being extended in physical space and time, and that a common 'thisness' may underlie extreme dissimilarity or distinct thisnesses may underlie great resemblance, I call *Haecceitism*.[152]

A defender of the naive counterfactual analysis of causation, who was also a haecceitist in something like Kaplan's sense, would presumably be unmoved by the case represented in Diagram 5. It represents a challenge to that analysis, only on the assumption that a distinct thisness may not underlie great resemblance. Such a haecceitist could say that the naive analysis rightly tells us that

---

[151] Lewis, "Postscript E", p. 195, footnote 20

[152] David Kaplan, (1975) pp. 722-723

without the firing of c_1, the firing *e* would not have occurred, even though something indistinguishable from it in time and manner would have occurred.

This may seem to suggest that the very demand for criteria of trans-world identity is as confused as the pupil in Kripke's discussion of a probability experiment involving the throwing of two six sided dice:

> Nor should any school pupil receive high marks for the question 'How do we know, in the state where die A is six and die B is five, whether it is die A or die B which is six? Don't we need a "criterion of trannstate identity" to identify the die with a six - not the die with a five - with our die A? The answer is, of course, that the state (die A, 6: die B, 5) is *given* as such (and distinguished from the state (die B, 6: die A, 5)). The demand for some further 'criterion of transstate identity is so confused that no competent schoolchild would be so perversely philosophical as to make it. The 'possibilities' simply are not given qualitatively ... If they had been, there would have been just twenty-one distinct possibilities, not thirty six.[153]

---

[153] Saul Kripke, Preface to *Naming and Necessity*, pp. 16-17. I agree with Lewis that the pupil deserves high marks despite her error. See *On the Plurality of Worlds*, p. 226.

Suppose the actual outcome of the probability experiment is the event (die A, 6: die B, 5).[154] We do not need to suppose that there would have been a difference in time or manner between this event and the counterfactual event (die B, 6: die A, 5), in order for us to distinguish them: why suppose that we need such a difference in order to distinguish the actual firing of neuron *e* in Diagram 5 from the one which would have occurred had c_1 failed to fire?[155]

Suppose we accept that we can draw a numerical distinction between actual and counterfactual events which are indistinguishable in time and manner. This position would allow the naive counterfactual analysis to agree with commonsense that the firing of c_1 is a cause of the firing of *e*, even in cases in which *e* would have fired at exactly the same time and in exactly the same manner without c_1's firing. This position makes it more difficult,

---

[154] For what I hope are obvious reasons I am following orthodox probability theory in using the term 'event' to describe the outcome of a probability experiment, rather than Kripke's 'state'. Of course (die A, 6; die B, 5) is not the only event which may be called 'the outcome' of this experiment; there is also, for example, the more robust event (die A, even; die B, odd).

[155] It should be clear to those familiar with this example, that my purpose in discussing it is different from that of Kripke.

however, to justify agreement with commonsense that the firing of c_2 is not a cause of the firing of e. How, that is, can we justify distinguishing between the actual firing of e, and the firing of e which would have occurred without the firing of c_1, but not distinguishing between the actual firing of e, and the firing of e *which would have occurred without the firing of* c_2?

A haecceitist could say that the latter pair of events have a common thisness, which the former pair lack, and refuse to defend the position with any criteria of trans-world identity; viewing the request for such criteria as symptomatic of a confusion similar to that of Kripke's 'perversely philosophical pupil'. But this would be merely refusing to answer the question. Of course, it's always possible to refuse to defend what one says, but it would be nice if we did not have to. Kripke's example does not show, as he seems to think, that the demand for criteria of trans-world identity is confused. What it does suggest is that such criteria need not be restricted to the time or manner of the events in question.

A better approach is suggested by Kaplan's analogy between the haecceitist attitude to how things (in this case events) are spread out in logical space, and how we view things as being spread

out in physical space and time. One could say that the latter pair of events (the actual firing of *e*, the firing of *e* without the firing of c_2) have a common history, which the former pair (the actual firing of *e*, the firing of *e* without the firing of c_1) lack.[156] The actual firing of *e*, like the one which would have occurred without the firing of c_2 and unlike the one which would have occurred without the firing of c_1, received its energy from a neuron which in turn received its energy from c_1.

Following the lead of my (attempted?) dissolution of other preemption problems we could adopt the position that every event is implied by an event which essentially occurs at the time it does, essentially occurs in the manner it does, and essentially receives its energy from its actual sources (including indirect ones). Such an event would occur in the region in which *e* fires, and would be a candidate for being 'the firing of *e*'. This would give us a sense in which the firing of *e* is counterfactually dependent on the firing of

---

156 This position is a form of *haecceitistim*, if we construe that doctrine to be the denial that trans-world identity supervenes on the qualities (*e.g.*, time and manner) of the entity in question. This is not the way Lewis construes the doctrine, see Lewis, (1986) pp. 220-248.

c_1, but not on the firing of c_2; consequently the naive counterfactual analysis can truthfully say that c_1's firing is a cause of e's firing, whereas c_2's firing is not.

I suspect that Lewiseans will be inclined to reject this approach, because the postulated event is (partly) extrinsic. Lewis and Maudlin have argued that overly extrinsic events should be treated with suspicion, because they appear to stand in relations of noncausal counterfactual dependence to genuine events which are distinct from them.[157] If Oedipus had not killed Laius, Jocasta would not have been widowed; but does that mean that the killing caused the widowhood?

> We do not want to waste our time on such *scheinprobleme* as exactly when Jocasta became a widow (instantaneously in some preferred reference frame? when she intersects the future light cone of the murder?) ... If this is superluminal causation, it is not the sort to be of any concern.[158]

Maudlin suggests that the way to avoid such *scheinprobleme* is to restrict the analysis to "local physical events":

---

[157] David Lewis, "Events", pp. 263-264

[158] Tim Maudlin, (1994) p. 128

> The notion of a local physical event need not be fully explicated, but locally observable changes, such as getting paler, setting off a photomultiplier tube, and rotating a filter, certainly count; becoming a widow certainly does not.[159]

If I understand Maudlin correctly, his position is that we can avoid worrying about whether or not the killing superluminally causes the widowhood by denying that it causes it at all (at least in the sense of 'cause' that 'we' are concerned with), and that we can do that by denying that the widowhood is an event (at least in the sense of 'event' that 'we' are concerned with), because of its extrinsic essence.[160]

We need not be persuaded by this reasoning. We could instead say that the killing does not cause the widowhood, because they are not distinct, rather than because the latter is not a 'local physical event'. In particular, we could say that the former event implies the latter event; not in the sense defined by Lewis, since they occur in

---

[159] ibid.

[160] Like Maudlin, I am assuming that Jocasta is present in the region of her widowing, *i.e.*, it occurs in Thebes, not at the crossroads. I do not think this is unambiguously true.

different regions, but simply in the sense that the occurrence of the former implies the occurrence of the latter.[161]

Lewis has recently presented a similar argument against "specifying events too extrinsically":

> ... it won't do to say that events a third of a century ago caused me to speak this sentence in the place where once I was a student. (Though those events did cause my speaking *simpliciter*.)[162]

Whether this will do or not seems to me to depend on what the events of a third of a century ago are; if the events are Lewis's being a student there, or Lewis's becoming a student there, I agree. It seems to me, however, that the reason that it won't do is not because of anything wrong with the extrinsic event - Lewis's speaking certain words in a place where he once was a student - but because the relations between this event and *those* events of a third of a century earlier are logical rather than causal. The later event implies the earlier events. Lewis suggests in "Events" that he would

---

[161] In other words the killing is inseperable from the widowhood, in the first, but not the second, of the two senses discussed in Chapter Five.

[162] "Causation as Influence", given at the Whitehead Lectures, Harvard University, March 11 and 12, 1999. Lewis once was a student at Harvard.

reject this position; claiming that these events are logically independent of each other.[163]

Rather than pursue this issue I think it is worth pointing out that Maudlin's and Lewis's argument, as Lewis recognizes, should not be construed as ruling out events which are partly extrinsic.[164] In fact, Lewis explicitly considers and rejects the doctrine that genuine events must be entirely intrinsic:

> For if an event were an entirely intrinsic property of regions, then it would have to occur in any duplicate of any region in which it occurs. ... It may be hard for big regions to be exactly alike in what goes on in them; but remember that submicroscopic events go on in submicroscopic regions, and electron-sized regions will have less opportunity to differ than larger regions do. We don't want a collision of two electrons to be one event that occurs in all the regions of the same world where two electrons collide in just that way; rather, we want to distinguish different collisions that occur once each.

---

[163] "Events", p. 264. In fact the two events Lewis discusses are the death of Socrates and the widowing of Xanthippe. It seems to me that the former event implies the latter event. Lewis's claim that they are logically independent of each other may be based on the conditions of occurrence he ascribes to the widowing of Xanthippe, which do not include the identity of the husband who dies. This would be true, if the death of Socrates essentially involves Socrates, but Lewis explicitly avoids committing himself to the "existence of events which essentially involve Socrates". (ibid.)

[164] Lewis takes this argument to rule out events which are "predominantly" extrinsic.

> If the regions do not differ intrinsically, then an event that occurs in one but not all of them cannot be entirely intrinsic.[165]

In fact, many of the entities which we would ordinarily consider to be events, and ordinarily think of as standing in causal relations, seem to be partly extrinsic, for example the discovery of Pluto.[166]

Suppose that someone remains suspicious of extrinsic events, or is convinced that the firing of *e* (or death of Edgar) that he or she had in mind was not even partly extrinsic: or suppose that she rejects the appeal to an event which essentially receives its energy from its actual sources, not because there is anything wrong with such an event, but because its relation to the preempting cause appears to be logical rather than causal. There is another approach to these cases of early preemption which has been suggested by Lewis himself, who indicates in one passage that they are too far-fetched to represent a serious challenge to the counterfactual analysis.

While considering whether it is plausible to eliminate redundant causation altogether by appealing to the fragility of the

---

[165] "Events", p. 264

[166] I borrow the example from Hartry field.

effect (*i.e.*, the common suggestion) he rejects one possible objection:

> Suppose we did follow this strategy wherever we could. Wouldn't we still have residual cases of redundancy, in which it makes *absolutely* no difference to the effect whether both of the redundant causes occur or only one? Maybe so; but probably those residual cases would be mere possibilities, far-fetched and contrary to the ways of this world. Then we could happily leave them as spoils to the victor. For we could plausibly suggest that commonsense is misled: its habits of thought are formed by a world where every little thing that happens spreads its little traces far and wide, and nothing that happens thereafter is quite the same as it would have been after a different past.[167]

If this were right, then we could assume that the absence of the 'preempting cause' would have made *some* difference to the effect, even if we cannot say what that difference would have been.[168]

Nonetheless I hesitate to accept that we can be confident that the 'preempting cause' will always make some difference to the effect. It is certainly easy to imagine the actual and counterfactual effects being simultaneous duplicates of each other. Although we

---

[167] ibid. pp. 197-198

[168] Unlike late preemption where we can always appeal to the time difference.

may be confident that such cases are unusual, we cannot be confident that they are 'mere possibilities': they may well be actual.

But unusualness may be enough to justify the position under consideration. It is possible that any inclination we have to think of the preempting cause as a genuine cause and the preempted alternative as not being a cause, is a mistake. This mistake could be the product of a persistent, though (in such cases) erroneous, assumption that 'the effect' would have occurred in a somewhat different manner, if not for its 'preempting cause'.

I hope that this position will grow on you. The unusualness of the actual and counterfactual events being simultaneous exact duplicates of each other is particularly clear in my macroscopic example, because of the size of the region in which the funeral occurs.

It is important to secure a sense in which Charley's decision is, and Cecil's decision is not, a cause of Edgar's death.[169] The naive analysis accommodates this need by appealing to fragile (though not too fragile) candidates for being that death. But, I submit, it is not

---

[169] After all, Charley, unlike Cecil, could be held legally responsible for Edgar's death.

nearly so clear that we need a sense in which Charley's decision is, and Cecil's decision is not, a cause of Edgar's funeral, if the funeral would have occurred in exactly the same way without Charley's decision. It is, I think, plausible to dismiss intuitions to the contrary as being attributable to the strangeness of the case, along with its similarity to more familiar cases; *i.e.*, cases of early preemption in which the absence of 'the preempting cause' would have made some difference to the effect.

I think there is no compelling reason not to adopt the same approach to Diagram 5. I submit that our causal judgments about neuron diagrams are not very clear. Given my claim that causal judgments are partly judgments about importance, and we have not been told why we should be interested in the causes of the neuron firings in question, this is only to be expected.

Chapter 8

Trumping Preemption

All of the putative cases of preemption I have considered are cases, in Jonathan Schaffer's terminology, of "cutting". That is, they are cases in which the preempted process does not go to completion, either because of the continuation of the main process beyond the effect (*i.e.*, late preemption), or as a result of a branch process that diverges from the main process before the effect (*i.e.*, early preemption). Schaffer offers a putative case of preemption which is not like this:

> Imagine that it is a law of magic that first spells cast on a given day match the enchantment that midnight. Suppose that at noon Merlin casts a spell (the first that day) to turn the prince into a frog, that at 6pm Morgana casts a spell (the only other that day) to turn the prince into a frog, and that at midnight the prince turns into a frog. Clearly, Merlin's (first that day) spell counts as a cause of the prince's metamorphosis and Morgana's doesn't, because the laws say that only first spells have such consequences. Nevertheless, there is no counterfactual dependence of the prince's metamorphosis on Merlin's spell because Morgana's spell is a dependency-breaking backup. Further, there is neither a failure of intermediary events along the Morgana process (we may dramatize this by stipulating that spells work

> directly, without intermediaries), nor any would-be difference in time or manner of the effect absent Merlin's spell, and thus nothing remains by which extant counterfactual accounts of causation might distinguish Merlin's spell from Morgana's in causal status.[170]

The presence or absence of cutting is irrelevant to my approach to alleged examples of preemption, although it is not to some others.[171] Consequently I will discuss the example, not *qua* putative case of noncutting preemption, but *qua* putative case of preemption which we can not plausibly dismiss either by appeal to the fragility of the effect, or by claiming that it is really a straightforward case of symmetrical redundancy.

It is not plausible to view this as an unambiguous case of symmetrical redundancy, because of the 'law' which guarantees that only the first spell will be efficacious:

> Had Merlin's spell been different, such as a spell to turn the prince into a goat, then the prince's fate would have been different: he would have become a goat ... Not so for

---

[170] Jonathan Schaffer, "Trumping Preemption: A New Counterexample to Counterfactual Accounts of Causation", *Unpublished Manuscript*, p. 1. Henceforth, "Trumping Preemption".

[171] See, for example, Murali Ramachandran, (1997) p. 273, "... in all genuine cases of causal pre-emption, ... the pre-empted processes do not run their full course ... All genuine causes, on the other hand, *do* seem to run their full course; indeed, they presumably count as genuine precisely because they do."

> Morgana's spell; had she cast a spell to turn the prince into a goat, to protect his humanity, or event to make the sky the color of gold, come midnight all that would happen would be the prince becoming a frog ... To call her spell a cause would be to countenance causes whose alleged effects were counterfactually oblivous to them.[172]

Although I cannot appeal to the fragility of the effect to respect the *intuition* that *Merlin's spell is, and Morgana's spell is not, a cause*, I think I can do so by appealing to the fragility of the cause, *i.e.*, the fragility of Merlin's spell.

If I am right about the profligate theory, two events occur in the region in which Merlin casts his spell (both of which are, or may be, candidates for being Merlin's spell); the more robust one is essentially a spell, but only accidentally a prince-into-frog-spell, the more fragile one is essentially a prince-into-frog-spell. The robust version of Merlin's spell is not a cause of the prince's metamorphosis into a frog, since to suppose it not have occurred is to suppose that Merlin cast no spell at all, in which case Morgana's spell would have done the job.

But, I claim, the fragile version of Merlin's spell is a cause of the prince's metamporphosis into a frog. That is, if the fragile

---

[172] "Trumping Preemption", p. 4

version of Merlin's spell had not occurred, the prince would not have turned into a frog. This seems to follow from the relatively uncontroversial possible world semantics for counterfactuals, according to which counterfactuals are true iff the 'nearest' possible world in which the antecedent is true is also a world in which the consequent is true. The nearest possible world in which the fragile version of Merlin's spell does not occur is not one in which he casts no spell, but one in which he casts a different spell. In such a world the prince would not have been turned into a frog, he would have been turned into something else. This gives us a sense in which Merlin's spell is a cause of the prince's change. No matter how fragile Morgana's spell is, however, it does not cause the prince's change, since it would have been the same whether she had cast a different spell or no spell at all.

This approach was suggested to me by a passage in which Lewis argues that not only do Fragile and Robust have different causes, they also have different effects:[173]

---

[173] Keep in mind that in this example Lewis's "weak event" is what I have called "Robust", and his "strong event" is what I have called "Fragile". I am sorry, if this is confusing.

> The first event (the weak one) caused Fred to greet John in return. The second one (the strong one) didn't. If the second one had not occurred - if John hadn't said "Hello" so loudly - the first one still might have, in which case Fred still would have returned John's greeting.[174]

Similarly, I reason that the fragile and robust versions of Merlin's spell have different effects. The robust one does not cause the prince to turn into a frog: the fragile one does. If the fragile one had not occurred - if Merlin had not cast a prince-into-frog spell - the robust one might still have, in which case the prince would not have turned into a frog, but something else.

Objection 1:

The closest world (to the world of the example) in which any candidate for being Merlin's spell does not occur is a world in which Merlin casts no spell, not one in which he casts a different spell. This is because worlds in which Merlin casts no spell are, like the world of the example, worlds in which the prince turns into a frog.

Reply:

The objection assumes, and I am inclined to agree, that closeness has something to do with similarity. Similarity, after the time of the antecedent, however, should play little or no role in our

---

[174] "Events", p. 255

calculations. Otherwise we would have to agree that if Nixon had pressed the button, nuclear war would have been avoided.[175] I am inclined to agree with Jonathan Bennett that similarity at the time of the antecedent is the only relevant similarity. Surely, all else being equal, a world in which Merlin casts a prince-into-toad-spell is, at that time, more like a world in which he casts a prince-into-frog-spell, than to a world in which he casts no spell at all. This seems to be the same reasoning that underlies Lewis's claim that the fragile version of John's greeting does not cause Fred's reply.

Objection 2:

Suppose Merlin very much wanted the prince to become a frog, but hesitated to cast this spell out of a general reluctance to cast spells.[176] Arguably, in this situation the nearest possible world in which the fragile version of Merlin's spell does not occur is one in which he casts no spell, since it appears that the change required in order for there to be a different spell, in this case, would be greater than the change required for there to be no spell.

---

[175] See David Lewis, "Counterfactual Dependence and Time's Arrow".

[176] Hartry Field has pointed out this possibility to me.

Reply

In this situation the relevant version of Merlin's spell - the one that causes the change - would be, not only essentially a prince-into-frog spell, but also essentially one which was performed by someone with a strong desire for this prince to be a frog accompanied with a general reluctance to cast spells. If it had not not occurred, a sufficiently fragile version of the effect would not have occurred either.

Since my approach to Schaffer's example has nothing to do with the difference between trumping and cutting, it would not be surprising if it helped with some cases of cutting also. Lewis has told a story which may be a case of cutting or of trumping: either way I think it can be resolved by appealing to the fragility of (candidates for being) the cause.[177] In this story soldiers who, in case of conflicting orders, will obey those of the superior officer, are simultaneously ordered to advance by a Sergeant and a Major. It seems natural to say that the Major's order causes the advance and the Sergeant's order does not. This may be a case of cutting, in which case it would be a case of early preemption in which the

---

[177] "Causation as Influence"

process leading from the Sergeant's order is stopped somewhere inside the soldiers' brains and/or minds. On the other hand, it may be a case of trumping.

Lewis has used this example to argue that trumping preemption should be taken seriously by those who object to taking seriously far-fetched scenarios involving magic. I want to use it to show that my strategy of appealing to the fragility of the cause can help in some cases of cutting. The naive counterfactual analysis tells us that a sufficiently fragile candidate for being the Major's order to advance (*i.e.*, one which is essentially an order by the Major to advance) is a cause of the advance, since if it had not occurred the Major would have issued another order and the soldiers would have obeyed it. This, I think, does justice to the intuition that the Major's order is a cause of the advance.

I think the significance of these examples has nothing to do with trumping. Rather it is the role various alternatives to the preempting cause play in our intuitive responses. In the earlier cases of putative preemption, alterations to the manner of the preempting cause would have made no difference to the effect (that we know of). In these cases, by contrast, it seems essential that

alterations of the preempting cause would lead to a different effect, whereas alterations to the preempted alternative would leave the effect the same. If I am right that every event is implied by an event which essentially occurs in the manner it does, an alteration of the manner of one event must be accompanied by the nonoccurrence of a more fragile event. This makes it natural to dissolve such cases by appealing to fragile candidates for being the cause, in a way which parallels my attempt to dissolve the earlier cases by appealing to fragile candidates for being the effect.

Conclusion

When discussing 'the common suggestion' I said that Lewis approaches questions about the essences of events as if their answers were in some way, and to some extent, up to us. This feature of Lewis's thought has been almost completely ignored, or misunderstood, in the literature provoked by his discussion of causation.

It first makes itself felt in his discussion of what he calls "a class of cases distinguished by doubt as to whether they exhibit redundancy at all"[178]:

> As it actually happens, the man dies on Tuesday morning, face down on the ground, his heart pierced by your bullet, with an entry wound in his back and an exit wound in his chest. Without your act he would have died on Wednesday evening, slumped in a chair, his heart pierced by someone else's bullet, with an entry wound in his chest and an exit would in his back.[179]

Lewis says that it is unclear whether or not this is an example of redundant causation. In other words, it is unclear whether or not the time or manner of the death that actually occurred is sufficiently

---

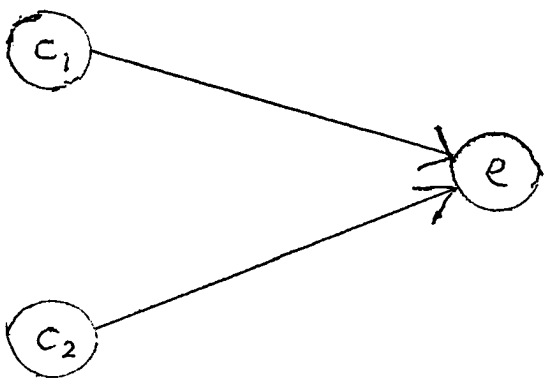
178 "Postscript E", p. 195

179 ibid.

similar to the one that would have occurred without your act to be the same death. Lewis also doubts that more detail about the actual and counterfactual situations would make the matter any clearer.

His second example of this class of cases is illustrated by the following neuron diagram:

Diagram 6:



In this case neuron *e* fires much more vigorously if it is stimulated by both neurons, than if it is stimulated by just one. Both neurons fire, doubly stimulating *e*:

> Is this vigorous firing of *e* a different event from the feeble firing that would have occurred if either one of c_1 and c_2 had fired alone? Then we have joint causation, in which the effect depends counterfactually on each of the causes, and there is no redundancy. Or is it that numerically the same firing would have occurred, despite a difference in manner, with single stimulation? Then we have redundant causation. Again it is hard to say, and

> again the difficulty cannot be blamed on underdescription of the details.[180]

Nor, I think, would Lewis say that the difficulty should be blamed on our epistemic faculties. It is not that we must resign ourselves to not knowing the answer to the question "Are these cases of redundant causation?": rather, Lewis is suggesting, that "there may be no right answer".[181]

Perhaps a better way to put the point, however, would be to say that there is no wrong answer, or at least that neither of the answers under consideration are wrong. I think Lewis would agree with me, that we can say that this is a case of redundant causation, or we can say it is a case of joint causation without redundancy, preferably on different occasions, and either way we will be right. By saying something, the truth of which depends on context, we can often create a context which makes it true.[182]

If I am right putative cases of preemption are also cases distinguished by doubt as to whether they exhibit redundancy or not.

---

[180] "Postscript E", p. 196

[181] ibid. p. 197

[182] I owe this insight to David Lewis, (1986) pp. 251-252, and "Scorekeeping in a Language Game".

In fact, if I am right there may be no realistic examples which unambiguously exhibit redundancy. The only difference between putative cases of preemption and the above two cases is that with the latter the ambiguity is between (symmetrical) overdetermination and joint causation, whereas the relevant ambiguity with the former cases is between (symmetrical) overdetermination and causation by the preempting, but not the preempted, cause.[183]

It is a tribute to the profound influence of David Lewis's work on the study of metaphysics in general, and causation in particular, that he has convinced many philosophers, including Tim Maudlin, Peter Menzies, Paul Horwich, Martin Bunzl, Douglas Ehring, Jonathan Bennett, William Goosens, David Armstrong, Michael McDermott, Bruce LeCatt, Richard Miller, Ned Hall, Jonathan Schaffer, Murali Ramachandran, Laurie Paul, Paul Noordhof, and others, that the most straightforward counterfactual analysis of causation breaks down in cases of preemption. The widespread acceptance of this dogma and, until now, the total absence of dissent, can partly be explained by

---

[183] As we saw there may be a legitimate (if nonstandard) sense in which putative cases of preemption exhibit joint causation.

the respect (and awe) which Lewis's work rightly receives from his fellow philosophers.

But the real explanation for what, if I am right, is a mistake, are the ways in which putative cases of preemption are represented - ways which make it natural to conflate different (though not distinct) events. Since neuron diagrams represent events as spatial regions, they cannot represent the difference between events which occur in the same spatiotemporal region. A similar conflation occurs when we describe these examples in natural language. I have argued that the event nominals of natural language are modally indeterminate - that is, they conflate events which occur in the same region, and differ only in their essential properties.

I certainly do not claim that I have proven that my position is correct. If anyone continues to believe, for whatever reason, that one or more cases of preemption are genuine, *i.e.*, they really do represent a situation in which an effect does not depend counterfactually on one of its causes, then I do not know how to convince them. But, I hope, I have at least undermined the popular notion that it is obvious that the naive counterfactual analysis of

causation is refuted by cases in which a causal proposition is true, even though its corresponding counterfactual is not.

Although my project has been largely destructive, I hope that something constructive (other than widespread acceptance of the profligate theory of events) may come out of it. I hope that it will allow us to focus on the real challenges facing the naive counterfactual analysis. I think that four such challenges remain; I will briefly discuss them in order of increasing urgency.

Firstly, my analysis should be accompanied by a satisfactory explanation of what *distinctness* is. I do not think that can be done in purely mereological terms, but I do not have any other suggestions.

Secondly, my analysis should have something to say about another kind of case which Lewis has argued demonstrates the possibility of a causal proposition being true, despite its corresponding counterfactual (as defined by the naive analysis) not being true:

> ... *c* occurs, *e* has some chance *x* of occurring, and as it happens *e* does occur; if *c* had not occurred, *e* would still have had some chance *y* of occurring, but only a very slight chance since *y* would have been very much less than *x*. We cannot quite say that without the cause, the

> effect would not have occurred; but we can say that without the cause, the effect would have been very much less probable that it actually was ... I think we should say that *e* depends causally on *c* and that *c* is a cause of *e*.[184]

Lewis proposes to analyse such 'probabilistic causal dependence' by means of a different kind of counterfactual; one with a consequent about the chance of the event in question not occurring. According to this analysis *e* depends causally on *c* (where *e* and *c* are distinct events) iff the probability of *e* (at a time immediately after *c*) is much greater than it would have been if *c* had not occurred. As before (*i.e.*, with the introduction of causal chains) we get from *causal dependence* to *causation* by taking the ancestral.

Putting aside the misgivings I have already expressed about Lewis's use of the causal-chain-modification, it seems to me that Lewis does not describe a situation in which we should say with confidence that *c* causes *e*. That is, I follow David Armstrong in suggesting that this is a situation in which we have the probability of causation, not a special kind of probabilistic causation. Lewis anticipates this position, and argues against it.[185] A detailed

---

[184] "Causation", p. 176

[185] ibid. pp. 180-181

response is beyond the scope of this dissertation, and quite possibly beyond my abilities.

Although I am often told that my analysis assumes determinism, I hope that this is not true, since I think it is quite possible that determinism is false. I do think, however, that we are often inclined to assume that determinism is true, especially when making causal judgments. Consequently, the discovery that our world was genuinely chancy may well show that some of our clearest (affirmative) causal intuitions are wrong.

Thirdly, since my analysis seems to entail that causation is not transitive, it should ideally be supplemented with a more persuasive argument against the widely accepted doctrine that it is. Neither Hartry Field's example which I briefly discussed, nor any of the other putative counterexamples to transitivity in the literature are conclusive. Each of them present three events, *c*, *d*, and *e*, accompanied by the apparently plausible claim that *c* causes *d*, *d* causes *e*, but *c* does not cause *e*.

There seem to be two possible responses available to anyone who, for whatever reason, remains attached to transitivity. In some cases it will be plausible to claim that the *d* which is an effect of *c*

is a different event from the *d* which is a cause of *e*. This approach, which seems to work with some of Michael McDermott's examples, is rather similar to my approach to preemption - both postulate an ambiguity between events which occur in the same region, to remove a threat to an analysis.

I chose to discuss Hartry Field's example, because I *cannot* think of any way to apply this approach to it, *i.e.*, I cannot think of two suitable candidates for being denoted by "my friend's stamping out of the fuse".[186] Another approach an advocate of transitivity can adopt is to claim that *c* is, after all, a cause of *e*, and try to explain away intuitions to the contrary.

Lewis has adopted this approach; suggesting that conflicting intuitions may be partly a result of the grip the idea that causation requires counterfactual dependence has on us.[187] He suggests that if, like everyone other than me, you have learnt to resist that grip in putative cases of preemption, you should not succumb to it in this case. I reply that the grip this idea has on us is what motivates me

---

[186] I do not rule out the possibility that this is attributable to a lack of imagination.

[187] "Causation as Influence"

to try to understand putative cases of preemption in a way which is consistent with it.

Fourthly, since my analysis entails that 'backtracking' and 'effects-of-a-common-cause' counterfactuals are, if true at all, true in a different sense from that in which 'causal' counterfactuals are, it should be accompanied by a semantics of counterfactuals which delivers this result. Lewis has presented a semantics for counterfactuals which does this.[188] Although I disagree with some of the details of this account, I have nothing better to offer.

---

[188] David Lewis "Counterfactual Dependence and Time's Arrow"

BIBLIOGRAPHY

Armstrong, David, M. *Universals and Scientific Realism Vol 2*, (Cambridge, 1978)

Armstrong, David, M., "The Open Door", in *Causation and Laws of Nature*, edited by Howard Sankey (Kluwer) forthcoming.

Austin, J. L., *How to Do Things with Words*, (Cambridge, 1962)

Bennett, Jonathan, "Counterfactuals and Temporal Direction", *The Philosophical Review*, (vol. 93, 1984) pp. 57-91.

Bennett, Jonathan, "Event Causation: The Counterfactual Analysis", *Philosophical Perspectives*, (vol. 1, 1987) pp. 367-386.

Bennett, Jonathan, *Events and Their Names*, (Hackett, 1988)

Bohm, David, *Causality and Chance in Modern Physics*, (Philadelphia, 1957)

Bunzl, Martin, "Causal Overdetermination", *Journal of Philosophy*, (vol. 76, 1979) pp. 134-150.

Bunzl, Martin, "Causal Preemption and Counterfactuals", *Philosophical Studies*, (vol. 37, 1980) pp. 115-124.

Byrne, Alex, and Hall, Ned, "Against the PCA-analysis", *Analysis*, (vol. 58, 1998) pp. 38-44.

Ehring, Douglas, "Preemption and Probabilistic Counterfactual Theory", *Philosophical Studies*, (vol. 56, 1989) pp. 307-313.

Ehring, Douglas, "Preemption, Direct Causation, and Identity" *Synthese*, (vol. 85, 1990) pp. 55-70.

Ehring, Douglas, "Preemption and Eells on Token Causation", *Philosophical Studies*, (vol. 74, 1994) pp. 39-50.

Ehring, Douglas, *Causation and Persistence: A theory of Causation*, (Oxford, 1997)

Ganeri, J., Noordhof, P., and Ramachandran, M, "Counterfactuals and Preemptive Causation", *Analysis*, (vol. 55, 1996) pp. 219-225.

Geach, Peter, *Reference and Generality*, 3rd. ed. (Ithaca, NY, 1980)

Goldman, Alvin, I., *A Theory of Human Action*, (New Jersey, 1970)

Goosens, William, K., "Causal Chains and Counterfactuals" *The Journal of Philosophy*, (vol. 76,1979) pp. 489-495.

Hall, Ned, "New Problems for an Analysis of Causation" *unpublished manuscript*

Hall, Ned, "Two Concepts of Causation" *unpublished manuscript*

Halpin, John, F., "The Miraculous Conception of Counterfactuals", *Philosophical Studies*, (vol. 63, 1991) pp. 271-290.

Hitchcock, Christopher, "The Role of Contrast in Causal and Explanatory Claims", *Synthese*, (vol. 107, 1996) pp. 395-419.

Horwich, Paul, *Asymmetries in Time*, (Massachusetts, 1989)

Hume, David, *Enquiries Concerning the Human Understanding and Concerning the Principles of Morals*, edited by L.A. Selby-Bigge (Oxford, 1962)

Jackson, Frank, "A Causal Theory of Counterfactuals", *Australasian Journal of Philosophy*, (vol. 55, 1977)

Kaplan, David, "How to Russell a Frege-Church", *Journal of Philosophy*, (vol. 72, 1975) pp. 716-729.

Kim, Jaegwon, "Causes and Counterfactuals", *Journal of Philosophy*, (vol. 70, 1973) pp. 570-572.

Kripke, Saul, *Naming and Necessity* (Oxford, 1980)

LeCatt, Bruce, "Censored Vision", *Australasian Journal of Philosophy*, (vol. 60, 1982) pp. 158-162.

Lewis, David, *Counterfactuals*, (Oxford, 1973)

Lewis, David, *Philosophical Papers Vol. 1*, (Oxford, 1983)

Lewis, David, "Scorekeeping in a Language Game", in *Philosophical Papers Vol. 1*. pp. 233-249.

Lewis, David, *On the Plurality of Worlds*, (Oxford, 1986)

Lewis, David, *Philosphical Papers Vol. 2*, (Oxford, 1986)

Lewis, David, "Counterfactuals Dependence and Time's Arrow", in *Philosophical Papers Vol. 2*. pp. 32-66.

Lewis, David, "Causation", in *Philosophical Papers Vol. 2*. pp. 159-213.

Lewis, David, "Events", in *Philosophical Papers Vol. 2*. pp. 241-270.

Lewis, David, *Parts of Classes*, (Oxford, 1991)

Lewis, David, "Many, but Almost One", pp. 28-44, in *Ontology, Causality and Mind: Essays in Honour of D. M. Armstrong*, edited by John Bacon, Keith Bacon, Lloyd Reinhardt (Cambridge, 1993)

Lewis, David, "Causation as Influence", *Whitehead Lectures at Harvard*, March 1999.

Lewis, David, "Void and Object", *unpublished manuscript*.

Loeb, Louis, E., "Causal Theories and Causal Overdetermination", *Journal of Philosophy*, (vol.71, 1974) pp. 525-544.

Lombard, L. B., "Causes, Enablers, and the Counterfactual Analysis", *Philosophical Studies*, (vol. 59, 1990) pp. 195-211.

McDermott, Michael, "Redundant Causation", *British Journal for the Philosophy of Science* (vol. 40, 1995) pp. 523-544.

McDermott, Michael, "Lewis on Causal Dependence" *Australasian Journal of Philosophy*, (vol. 73, 1995) pp. 129-139.

Mackie, Penelope, "Causing, Enabling, and Counterfactual Dependence", *Philosophical Studies*, (vol. 62, 1991) pp. 325-330.

Mackie, Penelope, "Causing, Delaying, and Hastening: Do Rains Cause Fires?", *Mind*, (vol. 101, 1992) pp. 483-500.

Maudlin, Tim, *Quantum Non-Locality and Relativity*, (Oxford, UK & Cambridge, USA, 1994)

Maudlin, Tim, "A Modest Proposal Concerning Laws, Counterfactuals and Explanations", *unpublished manuscript.*

Mellor, David, H., *The Facts of Causation*, (Routledge, 1995)

Menzies, Peter, "Probabilistic Causation and Causal Processes: A Critique of Lewis" *Philosophy of Science* (vol. 56, 1989) pp. 642-663

Menzies, Peter, "Probabilistic Causation and the Pre-emption Problem" *Mind*, (vol. 105, 1996) pp. 85-117.

Miller, Richard, W., *Fact and Method: Explanation, Confirmation and Reality in the Natural and the Social Sciences*, (Princeton, 1987)

Noordhof, Paul, "Problems for the M-set analysis of Causation", *Mind*, (vol. 107, 1998) pp. 457-463.

Paul, L. A., "Problems with Late Preemption", *Analysis*, (vol. 58, 1998) pp. 48-53.

Paul, L. A., "Keeping Track of the Time: Emending the Counterfactual Analysis of Causation", *Analysis*, (vol. 58, 1998) pp.191-198.

Paul, L. A., "Aspect Causation", *unpublished manuscript.*

Quine, W.V., *The Ways of Paradox and Other Essays: Revised and Enlarged Edition*, (Harvard, 1966)

Quine, W.V., *Philosophy of Logic*, (New Jersey, 1970)

Ramachandran, Murali, "A Counterfactual Analysis of Causation", *Mind*, (vol. 106,1997) pp. 263-277.

Ramachandran, Murali, "The M-set Analysis of Causation: Objections and Replies", *Mind*, (107,1998) pp. 457-463.

Rice, D. H., "Redundant Causation - Reply to David Lewis", *unpublished manuscript.*

Schaffer, Jonathan, "Trumping Preemption: A New Counterexample to Counterfactual Accounts of Causation", *unpublished manuscript.*

Schaffer, Jonathan, "The Varieties of Preemption: Further Counterexamples to Counterfactual Accounts of Causation", *unpublished manuscript.*

Smart, J. J. C., "Further Thoughts on the Identity Theory", *Monist* (vol. 56, 1972) pp.149-162.

Unger, Peter, "The Problem of the Many", *Midwest Studies in Philosophy*, (vol. 5, 1980) pp. 411-467.

van Fraassen, Bas C., "Singular Terms, Truth-Value Gaps, and Free Logic", *The Journal of Philosophy*, (vol. 63, 1966) pp. 481-495.

van Inwagen, Peter, "Ability and Reponsibility", *Philosophical Review* (vol. 87, 1978) pp. 201-224.

van Inwagen, Peter, *An Essay on Free Will*, (Oxford, 1983)