THEORIES OF CAUSATION

ZACHARY RAHMAN

UNIVERSITY COLLEGE LONDON

M.PHIL PHILOSOPHY

THESIS

APRIL 2001

ProQuest Number: U643801

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest.

ProQuest U643801

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT OF THESIS

I claim that causation is a basic feature of the world, of which no substantial philosophical theory can be given. I critically examine leading contemporary reductive theories that attempt to analyse or explain causation in terms of something else; in each case, I argue that the theory is inadequate in some fundamental way. Further, I suggest that such theories begin from an unduly restrictive viewpoint, and that once we broaden our purview to take in the range and heterogeneity of causal connections that we want to recognise in the world, causation appears not to be the kind of thing of which an illuminating theory should be expected. The thesis is set out in four main sections as follows :-

(I) _Deterministic theories of causation_

   I begin with some classic theories of causation whose forms and failings set the agenda for thinking about the topic. I introduce the problem of pre-emption, and its source; and raise sceptical questions about some traditional tenets including (1) the claim that causation must always be backed by laws, and (2) the idea that causation is in at least some sense not perceivable.

(II) _The Relata of causation_

   Here I take up an issue raised by (I): what kinds of things can be causes and effects? I argue that all the extant theories of causation give an unduly restrictive answer to this question, which is hard to justify. I propose a tolerant approach; there is no obvious reason why there should not be a _variety_ of kinds of causes or effects.

(III) _Probabilistic theories of causation_

   Most recent theories deny determinism and propose an analysis in terms of probabilistic relations. I show that none of these accounts work.

(IV) _Causation and Processes_

   A radically different proposal is considered: that causation might be identified with some empirically discoverable physical relation that connects causes and effects. Appealing to ideas previously raised, I show that this is not viable; its initial promise lies in the fact that it is less reductive - and to that extent closer to the truth - than the theories considered in (I-III).

TABLE OF CONTENTS

3

(I) DETERMINISTIC THEORIES OF CAUSATION

In this chapter, I will discuss two deterministic theories of causation: Mackie's
(`65) nomological theory, and Lewis's counterfactual analysis. Why, given the
fact that no-one currently holds either in the form in which it is presented here ?
First, they provide a good framework for introducing some important issues
about singular causation. What kinds of thing does it relate ? Must it always
involve laws ? Is it perceivable ? One's views on these questions constrain one's
account of causation and vice versa. Secondly, these theories are prototypes,
the core notions of which are passed onto their probabilistic descendants.
Correspondingly, they have a hereditary weakness: they are doomed to
counterexample because they fail to capture the idea, essential to our concept of
causation, that causation involves an *intrinsic* relation. This simple idea, which
I will call "Menzies second platitude" (after his 1996) will be spelt out more
fully in (IV).

I begin with Mackie's early theory, the essentials of which are laid out
below.

(1) The core analysis of a cause as an 'I.N.U.S.-condition' (i.e.) a present
condition that is in itself Insufficient for the effect, but is a Necessary part of
a present set of conditions that, jointly, is Unnecessary but Sufficient for
that effect. The key idea here is to maintain the treatment of causation in
terms of 'conditions', while accommodating the plausible claim that causes
need neither be necessary nor sufficient for their effects. Mackie's example:
it can be perfectly true that the short circuit caused the fire in my house
today, despite the facts (a) that a dropped cigarette would've done just as
well without the short circuit and (b) that the short circuit would not have
led to the fire had there been nothing around to burn.

It serves my purposes, and does no injustice to Mackie if we further simplify his analysis. First, we can broaden his talk of 'at least INUS-conditions', by dropping the 'I' and the 'U'. Causes *needn't* be insufficient for their effects; and when they are, the set of jointly sufficient conditions in which they are included need not be unnecessary, even if it usually is. So we have this: a cause is a necessary part of some total condition that is sufficient for the effect (i.e.) a cause is a condition sufficient in the circumstances for the effect.

(2) The explication of 'necessary condition' and 'sufficient condition'. These notions, fairly obviously, cannot be understood in terms of material implication. Mackie appeals to certain non-material conditionals that are to be interpreted in terms of 'deductive-nomological' form arguments that the asserter implicitly supposes to be available, in principle. For example, "A is a sufficient condition for B" is equivalent to the 'factual' conditional "since A occurred, B occurred", which itself reduces to the claim that there is a set of particular-fact premisses (S) and a set of law-premisses (L) such that, where (P) states that A occurred and (Q) states that B occurred, (L&P&S) implies Q.(Presumably, Mackie would add, in line with the traditional account of DN-explanation, that both P and L be non-redundant i.e. the implication fails if either is omitted.)

In sum, we might say that on Mackie's account, *nomic sufficiency* is the essence of causation.

Anscombe ('71 e.g., p.88) denies that this analysis is necessary for causation; she claims that it is mere philosophical prejudice that demands that laws must be involved. There are actually two aspects to her objection

(i) Does our concept require that causes *necessitate*, in some non-logical sense, their effects ?

(ii) Does our concept require that every particular cause/effect sequence instantiate some exceptionless generalisation ?

It is now widely accepted that the first of these is to be answered in the negative; that both scientific and ordinary thought allow for indeterministic causation. Anscombe gives a popular example in which I am supposed to wire up a

bomb to a geiger-counter and leave some uranium nearby. Maybe the uranium will decay in the period t, maybe it won't; it's a matter of chance. But if it *does*, then my depositing the uranium causes the explosion – even though we know that it didn't guarantee it.

Can we press this point further? Must a given causal sequence be an instance of *any kind of law at all* ? Anscombe might be seen as advocating just this kind of 'anomalism,' as opposed to mere indeterminism, which allows only that the laws supporting a particular causal sequence need not be deterministic laws. She claims that our notion of causing is best re-expressed in terms of 'arising out of', 'deriving from', or 'travelling from': "If A comes from B, this does not imply that every A-like thing comes from a B-like thing...or that every B-like thing has an A-like thing coming from it...or that given B, A had to come from it, or that given A there had to be a B for it to come from." (p.92) Might we not add: "or that B-like things are associated with some paricular probability of A-like things" ?

I think that Anscombe is right about our causal thinking; but a mere appeal to introspection – "just examine your concept of causation" – will not settle the matter. What other considerations might be appealed to?

Tooley (see e.g., Sosa and Tooley (`93) introduction p.13) thinks that there is an epistemological criterion: a nomological theory of causation can be correct only if causation is not "directly perceivable." He accepts the consequent, Anscombe denies it; this is what persuades him that she is advocating full-blown 'anomalism'. To put this in context, consider the standard genealogical account of such theories. The pre-Humean philosophical tradition assumed that there was some form of "must"- necessitation – linking causes to their effects; Hume (see e.g., his `75 sec.VII) called the tradition to account in the court of hardline empiricism, challenging us to locate the 'impression' of necessary connection in the perception of any given causal sequence that would distinguish it from a mere succession of objects. Confident that the challenge could not be met, Hume concluded that this connection must be found outside the particular sequence – in the repeated observation of similar sequences and/or the resulting subjective tendency to pre-emptively reconstruct such sequences in the imagination .

However, as commentators from Reid onwards have noted, not every observed uniformity will support a causal claim. Something more is required to confer causal status on the particular succession; hence the requirement that they be backed by laws (how much of a departure from Hume this represents depends on the metaphysical robustness of one's conception of laws). But now, notice that the notion of *determinism* is not essential to Hume's argument. Anscombe upbraids Hume for casting his philosophical vision in the wrong place: he should, she says, have rejected wholesale the philosopher's reconstruction of the concept of causation in terms of necessitation, rather than conceding it and merely questioning its empiricist credentials. However, Hume could have applied his reasoning in an exactly similar fashion to Anscombe's favoured notions ("arising from" etc.), challenging us to find an impression in the billiard-ball sequence corresponding to these. If Hume's argument is otherwise in good order, there is little point in faulting it on the grounds of i*gnoratio elenchi.*

But is the argument otherwise in good order? On the Humean view, we clarify a term by tracing its correlated 'idea' to its ultimate source in an impression (or impressions), either primary or secondary; any term whose lineage cannot be verified in this manner is to be abandoned as obfuscatory. As mentioned, Hume took it to be obvious that our notion of causation could not be grounded wholly in impressions of sensation. What this amounts to (given the aforementioned qualification about *necessitation*) is the claim that there is nothing relevant in our visual (say) experience that we can attend to or identify – in the way that we can attend to or identify material objects and their properties – that could be "the causal relation" – the "arising from"etc. Now, we really should not accept this as grounds for saying that we do not perceive causation. First notice that if we did, we would have a quick way with a whole host of relations; there is nothing special about 'the causal relation' in this. Admittedly, the fact that we cannot pick out (e.g.) the 'larger than'-relation in observing a mouse next to an elephant need not worry Hume. As such relations – to put it in contemporary terms – supervene on their relata (i.e: necessarily, if they exist then it holds between them) they are relevantly unlike causation. But there are plenty of nonsupervening or 'matter-of-

fact' relations which do not satisfy Hume's criterion of perceivability. For example, nothing one can reveal by considering my cup and this desk in mutual isolation will entail or imply that the former is *on* the latter. Nor can you visually identify the 'on-relation' in the way that you can identify the cup and the chair. However, this does not suggest that my being on the chair is not something that someone can visually perceive. In fact, I think we can perceive relations, including causation, in various sense modalities[1]. I will consider the visual case.

As I have just suggested, the fact that relations do not figure in our visual experience as Humean 'impressions' has no bearing on the question of whether or not we perceive them. Relations feature in the empirical world only in complexes, consisting of things-standing-in-relations: so if we perceive relations, we must perceive them as instantiated, as relating individuals. Now let us consider this in terms of the notion of perceptual content, or the way the world is presented to one in an experience, where 'ways' includes only what can potentially attended to in the experience[2]. It seeems perfectly obvious that we are presented in experiece with complexes involving relations; it would falsify the character of my current experience to deny that the spatial relations between the cup, the pen and my desk are an aspect of what is visually presented (and anyway, do I not *see* the spatial relations between the parts of each individual ?) Furthermore, our experieces seem to have causal content; we can say:

(i) Elizabeth saw the knife cut the butter

                    or

(ii) Michael saw the ball break the window

                    - and in each case, what is seen involves causation: X cuts or breaks Y only if X causes an appropriate change in Y. Does this not show that we see causation ?

---

[1] I think the case is strongest for tactual awareness. However, that case depends on arguments about the relation between bodily awareness and touch which would take me to far from my topic.

[2] this rider is a clumsy attempt to ameliorate a certain ambiguity in the term "way"(e.g.) the cup is presented to me as *red and on my left*; it is also presented to me *visually*. Both of the italicised phrasess could be 'ways'. The former picks out putative features of my environment, that figure in the content of my experience; the latter picks out a property of my experience that is not part of the content of that experience.

One might object as follows. Since issues about sense-data are not relevant here, we can agree that (a) if "A sees B" is true, then B is directly perceived[3]. However (b) expressions like "the knife cut the butter" and "the ball break the window" are manifestly not singular terms. Therefore (c) we cannot regard them as filling the argument-place in "A sees __", and so (d) (a) does not show that (i) and (ii) report direct perception of causation. This could be what Tooley (`90a, `90b) has in mind in his dismissal of Anscombe's point: "For the fact that something is observable in the ordinary, non-technical sense of that term does not imply that it is an object of immediate, or non-inferential perception..." (Sosa and Tooley `93 p.13)

Given the correctness of (b) and (c), we should ask: what *do* we directly perceive, according to (i) and (ii) and our assumption (a)? It does not seem that (ii) can be represented as

(ii´) Michael saw the ball at t and the ball broke the window at t

For (ii´) does not entail (ii) (imagine that one is only an inch tall and that one clings to the other side of the ball as it breaks the window.) Similarly, but more obviously, (iii) does not entail (iv):

(iii)Michael saw John at t and John killed Jim at t

(iv) Michael saw John kill Jim

Nor can we simply replace "see" in these cases with "see that", for (i), (ii) and (iv) have the characteristic extensionality of "sees": we can substitute coreferential expressions *salva veritate* in them[4]. I suggest that (i), (ii) and (iv) should be interpreted along Davidsonian lines[5], as involving an implicit quantification over a complex object of perception - perhaps an event. Then (ii) and (iv) would be represented roughly as:

(ii´´) ($\exists$x) Breaking (x) $\wedge$ Of (x, the window) $\wedge$ By (x, the ball) $\wedge$ At(x,t) $\wedge$ Michael saw x

(iv´) ($\exists$x) x is a killing $\wedge$ Of (x, Jim) $\wedge$ By(x, John)$\wedge$ At(x,t) $\wedge$ Michael saw x

---

[3] In fact, this would still need considerable refinement; see Snowdon `92.

[4] This needs qualification, at least to the extent that "sees" does not always allow existential generalization.

[5] For the general idea of Davidsonian analyses which appeal to covert quantification, see (II) and his `80e.

This would allow for the extensionality of (i), (ii) and (iv) alluded to above: if Jim is John's father, then both (iv) and (iv´) entail:

(v) Michael saw Jim's son kill John's father

This idea needs a lot more defence than I can give it here (see Higginbotham `83). All I am saying here is (1) that it is one way of elaborating the claim that we do talk, correctly, of perceiving causation and (2) that I do not see how to substantiate the claim that such perception is less than 'direct' or 'immediate'.

Anscombe, to recall, claims that causation is directly visually perceivable; Tooley claims that is only so in an uninteresting non-technical sense. Against Tooley, I would say that there is only one sense that we can put to Anscombe's thesis, viz. that the world is presented to us as one in which things act on, arise from, depend on one another. That is: our perceptual experience has 'causal content.' But nothing much follows about the nature of causation.

There are a number of choice points for theories of perceptual content. First, there is the 'conceptual vs. nonconceptual' issue: does having an experience, at t, with the content p entail that one is able, at t, to judge that p, or can the content of one's experience be correctly characterised in terms of concepts that one does not possess? Then there is the question of whether content in 'world-dependent' , in the sense that a veridical experience with the content p is such that it *could not* exist/be had in the absence of the state of affairs in virtue of which it is in fact correct. For my purposes, it will suffice to consider two possible combinations of these views. Suppose we decide that perceptual content is world-dependent and not purely conceptual. Then both the Mackiean/NeoHumean theorist and the Tooleian 'Causal Realist' (selected for their opposition) can accept that our experiences have causal content: I may have a experience as of one thing's causing another such that my experience could not be the way it is if the world were not the way it is. The only difference , I think, is that the neo-Humean will claim that what makes the state of affairs presented *causal* (ie, nomological backing) is not the same thing as that in virtue of which it is presented *as* causal. Tooley supports his epistemological criterion on the grounds that if a state of affairs

had its causal status partly in virtue of the laws instantiated in it, then we would be able to perceive causation only if we could perceive those laws; and we cannot. But that depends on the claim that a causal state of affairs can be presented as causal only if that in virtue of which it is causal is presented. The Humean need not accept that. One can consistently claim that laws are analytically prior to particular causal sequences while allowing that the particular sequences are epistemologically prior to the laws. Admittedly, this commits her to an 'error-theory' to the extent that , in perceiving a causal situation, we take the causal nature of the situation to be exhausted by what goes on in the immediate spatio-temporal locale; but presumably she was prepared for that. The same goes for a theory of perception that makes content world independent and wholly conceptual. Both the Neo-Humean and the Causal realist could accept, in this framework, that causal concepts were correctly deployed in veridical experience. This would leave open the question of what it is for something to be a causal state of affairs. So it does not seem that much ground can be made against a nomological theory of causation on such epistemological grounds.

I now turn to the sufficiency of Mackie's analysis, which is where I think the bigger problem lies. Suppose that we have more than one non-overlapping minimal sufficient condition present, and hence two or more INUS-conditions that do not belong to the same set of sufficient conditions, but only one causally relevant set of conditions. We can make this concrete by extending an example of Mackie's: two independent assassins fire on Brown within moments of each other, Smith's bullet doing the fatal job and preempting Jones's bullet by seconds. This is not a case of ('symmetrical') overdetermination; we have no problem in distinguishing Smith's shot as the cause of the death. But Mackie theory does. Letting C1 be a description of Smith's shot, C2 a description of Jones's shot and D a set of circumstantial premisses; we can suppose that there are laws L such that *both* (C1 $\wedge$ L $\wedge$ D) *and* (C2 $\wedge$ D $\wedge$ L) imply a statement E saying that Brown's death occurred. But it is not the case that both shots caused the death. Nor can we simply stipulate in our definition that the sufficient condition in which a cause is a

constituent must be exclusively sufficient, on pain of wrongly discounting Smith's shot as the cause.

The only kind of response I can see would claim that this in some way misrepresents the arguments that would be required to support the two causal conditionals **(i)** "since C1,E" and **(ii)** "since C2, E". However, there is a general reason to doubt that this can succeed, as the 2 cases are symmetrical in point of all the analytical tools that Mackie has at hand. For example, one might claim that in the expansion of (ii), the fact-premisses would have to be a set D′ that included the occurrence of Smith's shot, whereupon C2 would become redundant in the derivation, since (L∧D′) alone would imply E. But then, parity demands that the expansion of (i) feature a set D″ that records the occurrence of Jones's shot, resulting in Smith's shot being wrongly discounted as a cause.

It is easy to construct other cases that work in the same way, for instance with epiphenomena or joint effects. Suppose that **c** deterministically causes both **e1** and **e2,** the laws being such that neither could have come about except by being caused by **c.** We will be able to get from **e1** to **c**, and from **c** to **e2**, by DN-derivations; but neither of these derivations track *causal* paths. This confirms the naive view of causation expressed in Menzies' platitude. Laws, relating properties or types, are not located anywhere; they are too abstract to fully account for what happens here and now, when *this shot* causes *this death.* This, we naturally feel, is why Mackie fails to distinguish the real caus from the pseudo-cause. Is there room for a deeper diagnosis?

I now move on to Lewis's (1986b) counterfactual theory of causation (henceforward, "CTFC") whose motivating idea is that causes 'make a difference'. Lewis upholds 'Humean supervenience' as the core of his metaphysics; but, in a way, his theory of causation is reminiscent of the Berkeleian approach to material objects. It aims to capture what the concept of causation really *does* for us, without following the explicit letter of the vulgar view.

The substance of the theory is the sufficient condition for causation: counterfactual dependence. Where **e** and **f** are distinct, actually occurring events,

to say that **f** counterfactually depends (CFD's) on **e** is to say that if **e** had not occurred, then **f** would not have occurred; which is true iff there is a world in which neither e nor f occurs which is closer - (i.e.) more similar - to the actual world than any world in which **f** occurs without e (we do not assume that there is a unique closest world, or even that there are a number of worlds 'tied' for maximum closeness.)

The criteria for evaluating similarity amongst worlds are, in decreasing order of importance:

(1) Avoid major, ramifying violations of fundamental laws

(2) Maximise the spatiotemporal region of perfect match of particular fact

and

(3) Avoid small violations of law ('minor miracles')

- rough matching of particular fact is of little importance. These criteria are very important to Lewis, as we shall see.

CFD is nontransitive, and so too strong to be a necessary condition for causation; hence Lewis introduces the notion of a 'chain of stepwise counterfactual dependence' (a 'CFD-chain'), which holds between e and f iff there is a sequence of events <e1....en> each member of which CFD's on its predecessor, such that e1 CFD's on **e** and **f** CFD's on en. We can then say that e cause f iff there is a CFD-chain running from e to f.

We might notice at this stage that Lewis's theory represents a promising compromise between the intuitive view that causation is a relation that holds between particular relata regardless of whether it relates other pairs of the same type, and the nomological theory. On the one hand, the definition of causation explicitly mentions only particular events. On the other, since that definition reduces causation to counterfactuals, Lewis's semantics reduces counterfactuals to similarity relations amongst worlds, and his weighting system for similarity gives primacy to laws, the account is not a singularist one.

Lewis takes pains to compare and contrast his approach with that of Mackie, because, at first blush, it faces similar problems. Wherever we can construct a noncausal nomic sufficiency, a CFD-link is not far to seek. For instance, the joint

effects problem. Suppose **c** deterministically causes both **e1** and **e2**, and neither could have come about except by being caused by **c**. If e1 had not occurred, then given its actually determined history, we must suppose that c would not have been there to cause it, in which case **e2** would not have occurred. Therefore there is a CFD-chain from **e1** to **e2**, but **e1** does not cause **e2**. So Lewis's account must be wrong. And again, Menzies' platitude seems relevant. The counterfactual theory seems insufficiently sensitive to what goes on between e1 and e2 in making its causal judgements.

Lewis's response is as general as the form of objection. It appeals to the "no backtracking" ("NB") rule for evaluating counterfactual conditionals, roughly: in realising the antecedent, depart from actuality as late as possible, and check the consequent by letting things run on from there. *Don't* try and alter history to accommodate the change. So, the closest world where e1 does not occur is not a world in which c was not there to cause it, but rather a world in which c was there but somehow failed to bring it about. Since **c** still brings about **e2** in that world, **e2** fails to CFD on **e1** at our world.

In Lewis's pre-emption case, we have two independent systems of neurons (a1, a2 , a3) and (b1, b2, b3), hooked up in parallel to a synapse. Firings in the a-system are the events c1, c2, c3; firings in the b-system are the events f1, f2 ,f3 (the number corresponding to the neuron which is firing in the relevant system.), each system being set up so that the firing of the nth neuron in it stimulates the $n+1$th neuron to fire. The reaction of the synapse is an event **e**. We are to suppose that **a1** and **b1** both fire, inducing their subsequent neurons to fire. **c2** brings about **c3** which brings about **e**; but **c2** also inhibits **b2**, so that there is no actual **f3**. The problem, for the counterfactual theory is to show that **c1** but not **f1** causes **e**, which requires that it establish a CFD-chain on the (genuinely causal) 'a'-line only. Without NB, this cannot be done. For if the closest world in which **c3** did not occur is a world in which **c2** is not there to bring it about, then that is a world in which the b-line is not blocked, and so brings about **e**. That is, it is false that if **c3** had not occurred, e would not have occurred. But given NB, the closest world where **c3** does not occur

14

is a 'miracle'-world **w1** in which **c2** fails to bring about c3, but is still there to block **f2**. In w1, neither line runs to completion at **e**; therefore the counterfactual "if **c3** had not occurred, e would not have occurred" is preserved, and we can construct a CFD-chain from **c1**, but not from **f1**, to **e**.

Now this solution is highly contentious; furthermore, there is a less tractable kind of pre-emption (which Lewis calls 'Late' pre-emption, as opposed to the 'Early' kind above) for which it will not work. In this kind of case, there is no blocking of the pseudocausal / pre-empted line prior to the occurrence of the final effect: **c3** brings about **e**, which "doubles back" and blocks **f3**. In this case, the NB-principle doesn't help; since nothing on the a-line blocks the b-line, there is no event on the a-line such that if it had not occurred, the final effect would not have occurred. However, I am not going to say any more about these issues here, for the following reasons. Lewis devised his theory under the general assumption of determinism and, as he came to see, the possibility of indeterminism - which seems very likely to be actual - threatens his account. For suppose c indeterministically causes e, even though there is some slight independent probability of e occurring spontaneously without c. The crucial counterfactual fails; there are e-worlds (i.e. worlds in which e occurs) amongst the most similar not-c worlds. However, Lewis offers a probabilistic version of the counterfactual theory that can also tackle the deterministic cases; it is in the context of that theory that we must consider, in section (III), the pre-emption problems and his proposed solutions.

(II) THE RELATA OF CAUSATION

To ask after the relata of causation is to ask what kinds of entity are apt to be causes or effects, 'kinds' here meaning something like 'metaphysical categories.' Why is this a question worth asking?

All the theories of causation I will be examining bring with them quite severe restrictions on appropriate kinds. Mackie's causes and effects, loosely termed 'conditions' above are properties or *types*; Lewis's scheme demands (a quite specific conception of) *events*. Others appeal to *facts*. And those theories that posit some kind of causal process are constrained to give an account of causal relata that makes sense of how particular causes and effects link up with this process. However, our ordinary causal talk allows for a *multiplicity* of kinds of causal relata which is, I shall argue, hard to eliminate.

There are two ways in which a theory that demands a restricted range can respond to this. First, we could give an independent, prior justification for the restriction, by showing that the prima facie intransigent locutions can be reduced or explained away. Alternatively, we could settle for a post-factum justification: if the theory works well for the cases it *does* apply to, we should concede that the others are illusory. Generally, these strategies will work together: a more or less acceptable rejection of certain apparent forms of causal claim will be bolstered by the success of the theory that makes it.

As we will see later, none of these theories of causation work. This leaves the multiplicity of causal relata as an interesting datum about our concept of causation.

First, I need to indicate the heteromorphism and diversity to be found in our causal language. On the one hand, there are the relational expressions that can be used in making singular causal claims. Some of these are explicitly causal: "causes"(a dyadic predicate); "because" (a sentence functor); and "because of" (which stands between a

sentence and a singular term). Others include "as a consequence of"; "resulted in"("was a result of,"" as a result,,", etc.); "the effect of which was"; "due to"; "led to." On the other hand, there are the relatum-expressions: sentences, singular terms for 'substances'(including proper names and definite descriptions), and singular or quasi-singular terms introducing a variety of what we might call 'non-substantial things' -- events, tropes, and property-exemplifications. To add to this complexity, there are no obvious interesting restrictions on how these expressions can be combined in well-formed causal statements. Of course, grammar demands e.g., that "because" take a pair of sentences; but generally, there is considerable scope for interaction:

(1) Hitler caused the war

(2) The war happened because of Hitler

(3) John's refusal was due to the fact that he had found three rats in the kitchen

(4) His indignation at the Tories led to his polemic

(5) The coarseness of the towel caused the splotches

(6) The bridge collapsed because the bolt gave way so suddenly.

(7) The bolt's giving way so suddenly caused the bridge to collapse

We should also note the many implicitly causal statements, such as:

(8) John opened the door

(9) Mary broke the window

As noted in (I), causation is implicit in these transitive verbs -- the truth of (9) requires that Mary did something that caused the window to break. The form has been largely ignored in discussions of causation in favour of the explicit causal locutions, probably because of sheer numbers (dried, bent, squashed, hurt, rang ...): the causation s seems to be too 'dilute' in this form to aid analysis. Some, however, have been guided to significant conclusions by the all-pervasiveness of this idiom (Anscombe `71).

Our primary object is to discern the nature of the entities which claims like 1-9 cite as causal relata. "Hitler" seems unproblematic. "The explosion" and "the war" must

refer to events, if anything does. But we need to know some more about what events are: do they include *the bolt's giving way* and *his indignation at the Tories*? In (2) we have explicit mention of a fact; but to what category do the effects cited in the noun-infinitive constructions, (7) and (4), belong? Clearly, "the bridge to collapse" is not a singular term of any kind.

I propose to begin with a discussion of events. There are several reasons for this. The view that events are the sole relata of causation has been a widespread if not the received one in recent times; this is largely due to Davidson, who explicitly argues for it, and Lewis, whose counterfactual theory of causation requires it. Yet there is far less agreement on what events *are*. We need to evaluate the different conceptions put forward by Lewis, Kim and Davidson if we are to adjudicate the question of the ontological commitments of causal claims like 4, 5 and 7; and this is a good way of introducing the case for *facts* in causation.

Davidson has done the most to argue for events as a fundamental category of entities. He identifies five roles (see his `80b for a summary) for such entities to fill.

Firstly, he claims that any satisfactory theory of action must allow for literal talk of one and the same action under different descriptions. This requires us to posit a single entity -- a particular event -- that is being redescribed. We can explain your handing of the piece of paper to me only because your handing of a piece of paper to me *was* your handing of a ten-pound note to me, which *was* your repayment of the debt. Similarly for explanation in general. Explaining occurrences requires bringing them under laws or generalisations; but laws, as Davidson puts it, "are linguistic"; they cover events only as described. So we must suppose that there are real entities there to *be* appropriately described. For example, we can explain the disaster that occurred in Tucson last week, but only because the disaster *was* the forest fire: there are no generalisations framed in terms of 'disasters,' but there are valuable generalisations about forest fires. Thirdly, there is Davidson's famous argument for the identity theory of mind, which can be briefly summarised as follows. Mental occurrences stand in causal relations to physical ones; causal relations must be instances of strict laws; but there are no strict laws (again linguistic entities) involving mental (attitude)

predicates. Therefore the laws backing any given mental-physical interaction must be physical ones, and the mental event must be a physical one: identity. But then mental events must be bona fide individuals. Fourthly, Davidson's favourite argument posits event-individuals for the purposes of semantic theory. Consider:

(i) Jones buttered the toast

(ii) Jones buttered the toast with a knife

(iii) Jones buttered the toast in the bathroom

(iv) Jones buttered the toast with a knife in the bathroom

Each of (ii),(iii) and (iv) entail (i); but (ii) and (iii) together do not entail (iv). A semantic theory should explain these facts, ideally in terms of structure -- the situation would be essentially the same if we replaced 'buttered' with 'sliced.' If we treat i-iv as constructed from different predicates ('x① buttered x②'' , 'x① buttered x② with y ', 'x① buttered x② in z, etc.), we can secure the entailments only by specifying special rules of inference. But this seems implausible in a theory of actual English, not least because "buttered" has 'variable polyadicity': there's no in-principle upper limit on how it can be modified. Davidson's elegant solution is to represent i-iv as quantified event predications as follows:

(i′) [∃x: event x] buttering(x) ∧ Of (the toast,x) ∧ By(Jones,x)

(ii′) [∃x: event x] buttering(x) ∧ Of(the toast,x) ∧ By(Jones,x) ∧ With(the knife,x)

(iii′) [∃x: event x] buttering(x) ∧ Of(the toast,x) ∧ By(Jones,x) ∧ In(the bathroom,x)

(iv′) [∃x: event x] buttering(x) ∧ Of(the toast,x) ∧ By(Jones,x) ∧ With(the knife,x ∧ In(the bathroom,x)

In this way, the data are simply explained by the standard rules for conjunction and quantification. The basic method -- posit an underlying event -- has a wealth of applications that tell in its favour.

None of these arguments is insuperable; but I do not want to question the thesis that there *are* events. My purpose in reviewing Davidson's case has been primarily to lay down some essential background context for the rest of my discussion.

Firstly, it should now be easy to grasp the positive side of Davidson's proposal about causal relations. Demurring, as ever, from conceptual analysis, he merely offers a canonical logical form for causal statements: a dyadic predicate takes two singular terms, to yield an extensional context "a causes b." As we saw, Davidson accepts the neo-Humean thesis that any true causal statement -- say, "Mill's partaking of the oysters caused him to feel sick"-- must be backed by a law. But he also agrees with Ducasse (1926) that there need be no law involving "partakes of oysters" and "feels sick." It follows that the law covering that claim must be formulated in other terms. So, Mill's partaking of oysters and his feeling sick must be things that can be described in other terms, and since one *follows* and one *precedes*, these things must be events.

Second, it should now be clear that none of Davidson's arguments tell us much about what events are *like*. Take the argument from adverbs. The argument requires only that action sentences be interpreted as quantification over *something* to which the relevant predicates can properly apply. What kind of thing this is, is really quite underdetermined (as we shall see in discussing processes). Nor is this remedied in the current context by Davidson's minimal causal criterion for individuating events: e1 is e2 iff they have all the same causes and effects.

Kim (`66, `69, `76), by contrast, is far more interested in offering a metaphysical account of events than arguing for their existence. Kim's events are structured entities consisting of an n-tuple of objects *exemplifying* (this appears to be a primitive relation) an n-adic relation at a time (period or instant). A given event can thus be represented as [(x1...xn), Fn, t]. This is identical with [(y1...yn), Gn , t] iff t1=t2, F=G, and xi=yi. Thus, there is already plenty here for Davidson to disagree with. Davidson does not attribute any 'internal structure' to events (Kim calls Davidson's events 'structureless.') Nevertheless, it is not clear that any of Davidson's actual arguments for events precludes their having a Kimian nature. I will briefly mention a couple of obvious objections to this.

Davidson, as mentioned, takes events as fundamental entities i.e., metaphysically basic. Does this prevent him from accepting Kim's conception? If "metaphysically basic" means "not reducible to other metaphysical categories," then Kim's events are similarly basic, for [x, F, t] exists only if x exemplifies f at t, and cannot be reduced to the mere sum of its constituents. Of course, if "basic" means "has no distinguishable constituents," then Kim's events would not be basic. But Davidson really cannot claim this for his events; how could, say, a tennis match, fail to have any identifiable components (if only spatio-temporally 'smaller' events within it)?

It must be admitted that Davidson could never use Kim's events for his metaphysics of mind; for it is central to that picture both that (a) mental events are physical events and (b) physical properties cannot be identical with mental properties, and (given that mental events are exemplifications of mental properties) by Kim's criteria, (a) and (b) are inconsistent. However, this clash is not sourced from *within* Davidson's theory of events.

The real source of the conflict between them is the *semantic* reading of the Kimian identity condition, which Kim has held more or less explicitly through his (`66, `69, `76) (i.e.) that "[x, F, t]" and "[y, G, t′]" pick out (denote, describe) the same event iff "x" and "y" refer to the same object, "t" and "t′"refer to the same time, and 'F' and 'G' *denote the same property*. This puts a quite severe restriction on the possibility of redescribing events. Kim assumes, in this context, that there is, generally, a property

where there is a predicate (so that properties are 'abundant' in Lewis's `83a sense) and that two predicates denote the same property just in case the properties they pick out are the same by Lewis's criteria for individuating 'abundant properties' i.e., just in case they pick out the same class of particulars, actual and possible. So, for instance, if the same verb is differently modified in two event names, we will have two distinct events named. For Davidson, this is Mill's fatal error of "thinking that we have not specified the whole cause of an event when we have not wholly specified it....that every deletion from the description of an event represents something deleted from the event described."(`80a; p.157 in his `80e) Indeed, as we saw, redescription thesis was a central and intuitively plausible strut in his case for events.

Nevertheless, there is an equally intuitive case for Kim's view. Kim takes events to be entities that are in some sense, correlated with whole indicative sentences; he sometimes says that sentences 'report' or 'describe' events. Davidson has consistently opposed this view as wrongheaded and, on a sufficiently strong reading, bound for incoherence (see the 'slingshot' argument below). Singular terms refer, predicates describe: there is no application of these notions to whole sentences. But, we can see what Kim is getting at when we consider that both he and Davidson see gerundive nominalisations of sentences as the paradigm means for referring to events. Kim's "[x,F,t]" is just a slightly formal rendering of "x's f-ing at t", which is uncontentiously derived from the sentence "x F'd at t". Now Kim, reasonably enough, wants his events to have a role in causation and explanation. Consider the following case.

A certain rope is designed so that its fibres strengthen under gradual increases of tension. By slowly increasing the weight we suspend from it, we can get it to hold a ton; but it will break if a much smaller weight is sharply applied. It could be true to say

(a) the rope broke because it was pulled suddenly

but false to say

(b) the rope broke because it was pulled

Conclusion: "the rope was pulled suddenly" and "the rope was pulled" describe *different* events. If substituting one of these sentences for the other in a causal context can change the truth value, then how can the entities picked out by "the violent pulling of the rope" and "the pulling of the rope" not differ in their causal powers ?

However, this has apparently false consequences. We can think of "is a pulling" and "is a violent pulling" as expressing something like event sortals; they come with a principle for counting events of a certain kind. Call these E1 and E2 respectively.

(1)   We assumed that I pulled the rope only once; so E1 is uniquely instantiated in the context (hence we are able to talk about *the* pulling)

(2)   Similarly for E2 (hence *the* violent pulling)

(3)   Ex hypothesi, the pulling was violent (i.e.) the pulling was *a* violent pulling So, how can Kim avoid the conclusion that the pulling *was* the violent pulling?

By rejecting (1). Kim crucially distinguishes the constitutive property (or 'generic event', picked out by "F" in "[x,F,t]") which a given event is an exemplification of, from those properties merely exemplified by that event. What we assumed to be a single event conceals a myriad of Kimian events differing in their qualitative constitution. If e = [(me, the rope), pulling, t], then e *was* violent; but it was not constitutively so, although *another* event at the same location [(me, the rope), pulling violently, t] was. It is the constitutive property of an event that determines where it stands in the causal network.

In sum, then, Kim has a metaphysical thesis about events, and a semantic thesis about their names, which go together naturally but not necessarily. Kim could consistently hold that two intrinsic event-descriptions "x, F,t" and "x G,t" corefer even though 'F' and 'G' denote different properties; but he denies this because he wants his event ontology to reflect the kind of semantic data considered above.

Before saying more about the Kim/Davidson dispute, I want to introduce a third conception of events, devised by Lewis ('86b/c) and developed by Yablo ('92, '92b) which puts ideas similar to Kim's in the framework of Lewis's modal metaphysics.

Lewis's theory is particularly interesting because it custom-built for his worked out counterfactual theory of causation.

According to Lewis's modal realism, claims about how things might or must be are explicated in terms of how things are at other existent but non-actual worlds. Like Davidson, he takes events to be particulars, and not different from substances or individuals in any metaphysically interesting way. Lewis offers us two ways f thinking about particulars for modal purposes. Consider some actual object called "α". If we take "α" to refer to an entity that exists only at our world, then (since nothing exists at more than one world) we must interpret claims about how α might have been in terms of how α's non-actual counterparts are. Alternatively, we can treat "α" as referring to a "modal continuant": a transworld class of particulars. In that case, claims about how α might have been depend on how the *members of* α are.

On this second version, an event is a transworld class (i.e. Lewisian *property*) of spatio-temporal regions, no two existing in any one world. To be a bona fide event such a class must also satisfy some (vaguely specified) further constraints, the most important being that it must not be too "rich"- that is, the class must not be too restrictive. For suppose we decide to designate my typing of the word "designate" an event that could not have differed in the merest detail - a 1-member class. This event - call it "unit" - would be far too easily caused on the CFTC (almost every event x is such that if x hadn't happened, unit wouldn't have happened), and would cause far too little (there is no other event that wouldn't have happened if unit hadn't). This, says Lewis, is no event at all.

The motivation for this picture is reminiscent of Kim's examples, although Lewis rejects Kim's semantic thesis. Suppose that I say "hello" to you - rather loudly, as it happens - and you respond in kind. Ordinarily, we would simply say that my greeting caused your reply and think no more of it, and Lewis's theory agrees: if the former had not occurred, the latter would not have occurred. But on reflection, we would also agree that my greeting could have been different in many respects without its causal role altering. For instance, had my "hello" been a bit less loud, you would still have responded (In Kim's terms: you replied because I said "hello", but not because I

said "hello" so loudly). Call that fully concrete event that includes all the possible greetings I make to you at all the various possible worlds "e1"; let e2 be its richer companion that includes all and only the loud greetings. Let e3 be the complement of e2 relative to e1. Now, if e1 had not occurred, your reply would not have been forthcoming; this reflects the common-sense causal judgement above. But if e2 had not occurred, e3 would still have, and so you would still have replied. So, putting it in Lewis's terms, the "hello" that caused your reply was a loud "hello", but not an *essentially* loud "hello". The causal distinctions that Kim achieves by individuating events in terms of constitutive properties, Lewis achieves by means of essential properties.

Thus, as with Kim, we get a small ontological explosion. At the location of "the" greeting, we discover a plethora of "hello's", partially orderable by "richness", the members of each successive event being a subset of any earlier event's constituency. For any (not too relational) property F at the location, there is an event that is an essentially an F-greeting; for any collection, K, of such properties there is an event that is essentially a K-greeting; not to mention all those events that are essentially F but only accidentally greetings. (All these events are distinct but part-identical; since any event depends counterfactually on any less rich companions it may have, Lewis would be faced with a rebarbative explosion of intra-location causal relations if they were conceded to be wholly distinct).

Despite the greater complexity of Lewis's account, his events do not allow for as many causal discriminations as Kim's. Kim's kind of case would be one in which you react to me *only because* I say "hello" so loudly, or, putting it in Lewis's terms, where the richer greeting is required for the reply and the weaker greeting is not enough. Lewis's account is not sensitive to this difference because he treats the richer event (i.e. e2)as a subclass of the weaker event e1. This means that e1 could fail to occur only if e2 failed to occur also. Thus, any event that depends counterfactually on e2 - that is, is such that if e2 had not occurred, it would not have occurred - must also depend counterfactually on e1. So while Lewis can allow e1 to have *more* causal efficacy than e2, he cannot allow it to have less.

Yablo (92, 92b) resolves this problem in a similar framework by refining the counterfactuals conditions on causation. Our causal thinking, Yablo says, is governed by the principle that causes are commensurate with, or proportional to, their effects: they neither include too many irrelevant features nor omit too many relevant ones. He aim to capture this rule in four counterfactual conditions: contingency, adequacy, requirement, and enoughness. The adequacy condition is as follows:

[A] if c had not occurred, then *if it had*, e would have occurred.

To see how this odd sounding principle works, consider Kim's case in which you responded not because I said "hello", but because I said "hello" so loudly. What this amounts to, in Yablo's scheme, is that the greeting-event e2 that caused your response was essentially, rather than accidentally loud. There is no problem here. If e2 had not occurred, you would not have replied, and so e2 caused your response. Now we must concede that I could have said "hello" much less loudly; we must allow that there is an event e1 located in the same spatio-temporal zone that is only accidentally loud. The problem for Lewis was that if e1 had not occurred (that is, if I had not said "hello"), you would not have replied, so that the causal difference between e1 and e2 does not emerge. [A] solves this. Consider a world w where e1 does not occur. In the closest worlds to w in which e2 occurs (i.e. there is a loud greeting), you reply. But in the closest worlds in which e1 occurs, you do not reply. Why? Well, intuitively, since *loudness* is not essential to e1, the most economical way to imagine e2 into the world is as a 'mere' "hello", which will not, according to our story, earn a response. Hence e1 does not cause your reply.

Let us get our bearings. We have before us three more or less detailed conceptions of events, each motivated by plausible ideas, each claiming to be a characterisation of the sole metaphyical category from which causes and effects are drawn. There are notable similarities and disagreements amongst them. Davidson's and Kim's events are sui generis in a way that Lewis's (a species of class) are not. Kim is unique in explicitly conceding a role for *properties;* as I read him, his assignment of structure to events is importantly motivated by the thought that we naturally discern the

properties in virtue of which one thing causes another. Nevertheless, while Lewis is like Davidson in denying this, he thinks that we need a lot more events than Davidson would allow in order to account for the kinds of distinction that Kim makes.

To make a start on adjudicating this dispute, I want to assemble some linguistic data concerning sentence nominalisation (Vendler `57a; Chomsky 1975)

As mentioned above, our primary means of picking out events are 'sentence nominals': nounlike expressions that can be formed from whole sentences. We can distinguish two basic forms of nominal: the *perfect* and the *imperfect*; I indicate the two ways of getting a nominal out of a sentence below with "IN"(imperfect) and "PN" (perfect):

[S1] "The referee intervened in the third"   yields

[PN1] "The referee's intervention in the third"   and

[IN1] "The referee's intervening in the third"

[S2] "Caesar died on the Ides"   yields

[PN2] "Caesar's death on the Ides"   and

[IN2] "Caesar's dying on the Ides"

[S3] "John built the house"   yields

[PN3] "John's building of the house"   and

[IN3] "John's building the house"

Notice that the perfect nominal may be formed from a deverbative noun, as in case [1] (similarly: death, betrayal, destruction), or as in cases 2 and 3, from the gerund of the verb followed by a genitive "of".

While both forms of nominal are fit to stand as the subjects of sentences, the perfect nominal is more closely allied with the more familiar nouns; the verb-al element of the parent sentence remains 'alive and kicking' (Vendler op cit.) in the

imperfect. This is the key distinction between them, and the source of the systematic differences in their syntactic and semantic behaviour described below.

(i) The PN-form allows for singularisation (via the definite and indefinite articles) and pluralisation. Hence, "an intervention" and "the sudden snapping of the rope" are fine; "an intervening", or "a snapping suddenly" are deviant.

(ii) The PN-form will accept adjectival modification in attributive position ("the referee's stupid intervention"; "the sudden snapping of the rope"); the IN-form willingly accepts adverbials ("The referee's stupidly intervening"; "the rope's suddenly snapping")

(iii) The PN tends to fit in contexts of observation, location, and temporal structure (e.g.) "the referee's intervention occurred just before the bell" "we were all amazed to witness the intervention" make sense, as do "Caesar's death took five minutes", "I myself observed Caesar's death on the Ides". By contrast, " we were all amazed to witness the referee's intervening", "the referee's intervening occurred just before the bell", "Caesar's dying on the Ides was sudden and painful", "I myself observed Caesar's dying" are all improper. Further, perfect nominals associate naturally with demonstratives ("this sudden intervention is intolerable") in a way that imperfect nominals don't ("this suddenly intervening is intolerable").

(iv) The gerunds in imperfect nominals can be negated ("Caesar's not dying on the Ides", "the referee's not intervening in the third") and tensed ("Caesar's having died on the Ides"). Neither works for perfect nominals.

(v) Substituting one form for another in a single context can make a semantic difference, notably in causal contexts. The truth of "the referee's intervening in the third round amazed us" entails that we were amazed *that he intervened.* But the gangsters who have prearranged that the referee will intervene at that point can still truly say "the referee's intervention amazed us", even though they are not in the least amazed that he intervened. For there are many aspects of the intervention other than its being an intervention that may be potentially surprising - maybe the referee hopped in on one leg, wearing a dress. Similarly, if Caesar's dying caused

the people to riot, then there was a riot *because Caesar died.* The parallel entailment for the perfect nominal doesn't hold. For suppose that the local residents are so weary of being disturbed at night that the noisy to-do as Caesar is stabbed incites an outbreak of public disorder. Caesar's death, then, *caused the riot*; but the people rioted because they were sick of being disturbed, *not* because Caesar died. More generally, we can note that the imperfect nominal from can usually be re-expressed as a 'that-p'(or even 'fact that-p') nominal without loss; this can rarely be done with the perfect.

Now for some comments and consequences.

The commonsense scheme of what there is includes, alongside regions for 'substances' and their properties, a region for things that *obtain, happen or go on*, which substances enter into or undergo. This is where we locate events, states, processes etc. What I have been attempting to elucidate above are two (non-exhaustive) ways in which our language carves up this region. On the one hand, there seems to be a "concretising" or "particularising" way of thinking, in which we treat what goes on at some portion of space-time as an object, a multifaceted thing with a nature beyond what we know or say about it. On the other hand, there seems to be a way of hatching a structure of things and properties or relations which involves abstracting from the greater detail that the world always exhibits.

This distinction is roughly tracked by, although not perfectly mirrored in, the perfect/imperfect nominal distinction. The difference between the two ways of thinking is reflected in a complex of interacting factors, including the meanings of verbs, tense, adverbial versus adjectival modification, and pragmatic or even stylistic factors that embody different points of view that we can take on the world.

Consider a couple of examples. John visits the market, just once, on Monday, say. We have the two quasi-singular terms "John's visit to the market" and "John's visiting the market". The former concretises or particularises the world. John's visit seems thing-like, something with a life of its own, that we can make discoveries and say things about: *it* was sudden, unexpected, enjoyable. The latter abstracts from the

world. In using the word "abstract", I do not mean to suggest that it involves us in talking about some *other* world e.g. of sets or universals or whatever. We're still talking about the everyday, empirical world when we use an expression like "John's suddenly visiting the market yesterday". Nevertheless, we seem thereby to be introducing just John, suddenly visiting, the market - and *excluding* any other things or properties. We are expressing the concept of a state-of-affairs or fact.

Anyway, the plausible hypothesis that I think is supported by the foregoing is that the use of the PN-form generally expresses the Davidsonian conception of an event-particular, whereas the IN-form expresses the idea of a Kimian event, which I think is a fact or state-of-affairs (as do Mellor '91; Bennett '88). First consider the Davidsonian side.

(a) Suppose that the market is in Berwick St. John went to Berwick St. just once yesterday, we can talk of John's excursion to Berwick St. yesterday. Now John's visit to the market must have been *an* excursion to Berwick St; and then it is very hard to resist the identity claim: John's visit to the market was John's excursion to Berwick St.

(b) If he had visited the market several times yesterday, we could have counted them e.g. "John's third visit to the market yesterday." Given that John's third visit was sudden and unexpected, we can talk of John's sudden unexpected visit to the market, which *was* his third excursion to Berwick St.

(c) If I know nothing about Berwick St., I may be puzzled that John went to Berwick St. But if I can reason in accordance with (a), I can alleviate my puzzlement, for I am not at all puzzled that he went to the market. Nevertheless, by the same reasoning, it was John's visit to the market just as much as his excursion to Berwick St, that caused my puzzlement.

Now let's look at it from Kim's point of view.

(a′) We have seen, by the Davidsonian argument above, that if John ran to Berwick St just once yesterday it is hard to avoid the conclusion that

(i) John's run to Berwick St was John's making of his way to Berwick St on foot. Now, we may assume that

30

(ii) John was tired because he ran to Berwick St.

is true; even though

(iii) John was tired because he made his way to Berwick St. on foot

is false. Therefore the Kimian entity correlated with or described by "John ran to Berwick St." is not the same as that correlated with " John made his way to Berwick St. on foot"

(b′) Let's see what happens when we introduce imperfect nominals - Kim's favoured form of intrinsic event-name. We find this kind dispute between Kim and Davidson.

"Notice...that it is not at all absurd to say that [John's running to the market] is not the same as [John's making his way to Berwick St]. Further, to explain [John's running to the market (why John ran to the market)] is not the same as to explain [John's making his way to Berwick St. (why he made his way to Berwick St.)]" (Kim `66 p.232n)

" The plausibility in this is due to the fact that not all [makings-of-ones way] are [runnings]. But this does not show that this particular [making of his way] was not [a running]...it was in fact, although of course not necessarily, identical with [John's run to the market.]" (Davidson `80b p.171)

Switching from imperfect nominals (Kim) to perfect nominals (Davidson) makes a difference.

Notice that it is hard even to express the Kimian opposition. Although we don't often formulate explicit event-identity statement in terms of perfect nominals, they seem natural enough. By contrast, "John's making his way to the market was John's running to Berwick St." is hard to interpret; we need to hear it in terms of perfect nominals. Similarly, the questions "Was John's making his way to the market identical with his running to Berwick St ?" "Was Brutus's killing Caesar identical with Brutus's stabbing Caesar?" They're hard to interpret. Either we put them in terms of perfect nominals - "was my pulling of the rope identical with my jerking of the rope?" - or, as Kim does, we use the 'same as' locution; and notice

how it then becomes natural, almost unavoidable, to drop the tensing: "Is John's making his way to Berwick St. the same as John's running to the market?" "Is Brutus's killing Caesar the same as Brutus's stabbing Caesar?" And, as Kim says, it's not at all absurd to answer 'no' to these questions. This is because Kim's nominals, although nounlike, don't pick out particulars. Perfect nominals come with a counting-principle. "John's run to the market" embeds the event-sortal 'run to the market', which allows us to count runs-to-the-market; and likewise 'sudden snapping of the rope', 'killing of Caesar'. Imperfect nominals aren't like this, which is why they don't allow count adverbials; we can't talk of John's three goings to the market or John's third going to the market

In sum, we should say that John's run to the market was his making of his way to Berwick St - they're one and the same event - but that John's running to the market is not the same as John's making his way to Berwick St i.e., not the same fact. The distinction was implicit in the initial dispute concerning how sentences relate to events. Kim's 'events' are 'described by' or correlated with whole sentences; but surely that makes them facts, situations or states-of-affairs, if anything. My provisional conclusion is that Kim and Davidson are elucidating two different concepts, and there should be a dispute between them only if either (1) one or other kind of thing is supposed not to exist or (2) one or other kind of thing cannot serve as a basic kind of causal relatum. As we shall see, there are objections to both facts (Davidson) and events (Mellor) under (2).

What about the Lewis-Yablo (LY) accounts ? It appears at first to be a "best of both worlds"; but I don't think it works. Let me explain why.

Kim's insight - obvious to many - is that on thing causes another in virtue of some, but not all, of its properties. As we saw, Kim makes this plausible by considering strong intuitions about causal statements, for example

(i) The bridge collapsed because the bolt snapped

(ii) The bridge collapsed (not because the bolt snapped but) because the bolt snapped so suddenly

We can easily imagine a situation in which (i) is false but (ii) is true. Yablo (92, 92b) provides one. The bolts in this bridge are specially designed so that, as long as they deteriorate at less than some specified rate, a safety mechanism operates, shifting the weight the bridge away from the area of weakness and strengthening the overall structure.

The LY-account uses our natural intuitions about the place of properties (e.g. suddenness) in instances of causation to gerrymander event-essences that, in turn, ensure the truth of those counterfactual claims required in order to bring the deliverances of their theory of causation in line with our intuitive causal judgements - where the theory analyses causation in terms of the counter factual dependence between 'brute' events. But this seems back-to-front. If we are in the first instance sensitive to the causal role of particular properties - thus, we say: "if the snapping had not been sudden the bridge would not have collapsed" - why not leave it at that? Why obscure this apparent fact by insisting that causation must only relate events *simpliciter* ? Were it not for the fact that we begin with a good idea of when and which properties are involved in a given case of causation, we would be lost as to which of the numerous LY-events was doing the causing. The combined effect of the events-with-essences metaphysic and Yablo's refined counterfactual constraints is to simulate the role of facts - i.e. what are reported by true sentences and picked out by imperfect nominals - in causation.

I say that the bolt's snapping so suddenly caused the bridge to collapse; Lewis and Yablo say it was rather that event that was essentially a sudden snapping of the bolt. There's something worth disagreeing about here: in order to get the same results, the LY-account uses considerably more complicated machinery, and does more violence to ordinary ways of thinking and speaking by introducing serious indeterminacies concerning the objects of our familiar nominals. For Lewis and Yablo claim to disown Kim's semantics; "the bolt's snapping suddenly" and "the bolt's snapping" could *each* properly refer to *either* the accidentally sudden snapping or the essentially sudden snapping. But this leaves unexplained the fact that if - as we would ordinarily say - the suddenness of the snapping was irrelevant to the collapse (e.g. the bridge

would've collapsed, *however* the bolt snapped), "the bridge collapsed because the bolt snapped so suddenly" is false.

How solid is this 'fact'? Pretty solid, I think. Certainly the statement is intuitively false in such circumstances; and there is a good reason for this. "Because" as commonly noted in this connection, is an explanatory connective. It isn't always a *causal* one, but this is not due to ambiguity; it's just that explanatory relations aren't always supported by causal ones. Now properties are crucial to explanation; and *irrelevant* properties are fatal to an explanation, as Salmon puts it. To borrow an example from him, suppose that I immerse X in Y and it dissolves. We can explain this by citing the facts that X is salt and Y is water (and maybe a law as well); but not by citing the fact that X is *hexed* salt or the fact that Y is *holy* water. "Because Y is holy water" is not just a *bad* explanation of this fact, it's no explanation at all. As a result, "because" creates 'property-sensitive' contexts , as it were. In the causal case, if suddenness is causally irrelevant, it's explanatorily irrelevant, and " the bridge collapsed because the bolt snapped so suddenly" is false. Lewis and Yablo can't explain why.

The LY-story was supposed to combine what's right in Kim's metaphysics - property sensitive causal discriminations - with the plausibility of Davidson's redescription semantics, i.e. that modifying the description of an event isn't modifying the event described. The attempt fails, in my view, once we realise that Davidson and Kim are talking about different things. Facts allow for the 'grain' of causation, as revealed in Kim's examples; events allow for redescription. But you can't give a non-relational redescription of a fact. This is because intrinsic non-relational expressions for facts don't name them in the same way that intrinsic expressions for events do, but express them. The fact that the bolt snapped is not the same fact as the fact that the bolt snapped suddenly, because the latter introduces the property of suddenness; there's no question of the two expressions 'referring to' the same thing. With perfect nominals, it's a different matter.

Of course, that doesn't sink the Lewis theory of causation. For one, he and Yablo both have worked-out metaphysical schemes in the context of which their accounts of

events are not ad-hoc in the way that I have been presenting them. More importantly, as mentioned at the outset of this section, it's only the theory of causation that will decisively test the theory of the relata, not vice versa. If Lewis's theory of causation is successful, then we've got a reason to bite the bullet and accept some revisionary metaphysics about events.

What I have aimed to show so far is that it's hard to make a convincing case for an events-only ontology for causation; the claim adds a burden of proof onto the theory of causation. And that's only half the matter. For facts are not the only prima facie candidates for inclusion in a pluralistic inventory of causal relata.

However, before considering those other candidates, we must deal with Davidson's case against facts. I mentioned above that the fact-theorist and the event-theorist must clash only if one or other kind of entity is either (1) denied existence, or (2) claimed to be unsuitable to provide causal relata. Davidson uses the 'slingshot' as a 2- argument against facts.

From an abstract point of view, the slingshot aims to show that if a sentential operator or connective ϕ is extensional with respect to singular terms (i.e. allows substitution of coextensive singular terms in its operands) then it must be truth-functional : the truth value of $[\phi(S,S')]$ is fully determined by the truth-values of S and S'. Let ϕ be "__because__". If the argument works, then "p because q" depends only on the truth-values of p and q. Since we know that **(a)** "p because q" can be true only if both of p and q are true and **(b)** facts are supposed to be correlated with true sentences, this would mean that if "because" reports a causal relation between *any* pair of facts, it reports a causal relation between *every* pair of facts (including the reflexive pair <p,p>). Since this is obviously unacceptable, we have a reductio of the claim that "because" expresses a relation between facts.

The argument runs as follows. Take any true sentences p, q, r, s.

(1)   logically equivalent sentences can be substituted *salva veritate.*(assumption)

(2)   coextensive singular terms can be substituted *salva veritate* (assumption)

(3)   p because q

(4) $\{x \mid x = x \ \& \ p\} = \{x \mid x = x \}$ because q   (by 1)

(5) $\{x \mid x = x \ \& \ r\} = \{x \mid x = x \}$ because q   (by 2)

(6) r because q   (by 1)

(7) repeat (3-6) for q and s

(8) r because s

Given that statements of the form "S because S'" don't report causal relations, what *do* they do? Davidson suggests they are (mere) "rudimentary causal explanations". 'Because'" is to be read as "causally explains", where causally explaining is not the same as the causing expressed by the 2-place predicate 'caused' which relates events.

I have nothing original to say about the slingshot; I will merely relate a now familiar point, made by Barwise and Perry ('81, '83), to show that anyone who wants facts in causation (or anywhere else for that matter) can reject the slingshot with a clear conscience.

The objection is that assumption (2) is dubiously applied here. Barwise and Perry effectively accuse the slingshot of equivocation, as follows. The slingshot cannot work without substitution of definite descriptions. But there are two ways of interpreting definite descriptions. First, there is Russell's way. On this view, they are 'incomplete symbols'; when they stand as subject expressions in atomic sentences, their semantic function is not to introduce an object which the predicate characterises, but rather to assert the unique instantiation of a certain property. Then there is Donnellan's "referential usage", in which the description is assigned an object (either that entity which does uniquely instantiate the property, or as Donnellan suggests, that entity intended by the speaker in the context). Now, say Barwise and Perry, we need not take a stand on the correct treatment of descriptions; but we must insist that a single definite description given the same semantic treatment throughout the argument. For concreteness, let p = "the match lit", q = "the match was struck" and

r = "London is in England". Consider the step from (3) to (4). If the description on the LHS of (4) is given a Donnellanian interpretation, the step fails even given (1). For, on that interpretation, the LHS of (4) simply says of a certain object (a set) that it

is identical with itself. So, the argument cannot even get off the ground unless "{x | x = x & p}" is given a Russellian reading. Next consider the step from (4) to (5). Obviously this fails if the description in (4) is given a Russellian interpretation. On that view, 4 and 5 introduce complex quantifications involving quite different properties. So, we can get from 4 to 5 only by giving the description a Donnellanian reading. But, as we saw, that prevents the step from 3 to 4. Therefore we cannot get from 3 to 5 without an equivocation. So the argument fails.

What about Davidson's account of those prima facie causal statements that don't fit his canonical form? Well, there is no doubt that

(A) Jane's coughing so loudly caused the conductor to wince   and

(B) The conductor winced because she coughed so loudly

are explanatory in a way that "her cough caused the conductor to wince" is not; "because", as we have said, always indicates an explanatory relation. But Davidson must deny that this relation is causal in any way at all. Further, since "cause" can link entities other than events, as in (A), Davidson must claim a semantic ambiguity.

Neither of these claims is plausible. As Mellor points out, Davidson's severing of causal explanation and causation makes it very hard to see what it could be in virtue of which the causal (B) differs from the noncausal

(C)   Jane's goal didn't count because she was offside.

It's natural to think that causal explanations work because they cite causes. But Davidson can't allow this, nor can he explain how 'underlying' causal relations might hook up with true causal explanations. There is no case, I conclude, for an events only causal ontology.

I now turn to a rival position pressed by Menzies (`89) and Mellor (`91, `95): the restriction of causal relata to *facts* only.

Mellor analyses causal claims in terms of probabilistic counterfactual dependence between chances, and chances require propositions or facts to apply to. He accepts that there are causal truths of the form "c causes e", where "c" and "e" refer to particular concrete events; but he claims that they are always derivative on factive causal claims of the form E because C (or: " the fact that C caused it to be the case that E"). Mellor argues that (i) causation can't relate *only* events, because there aren't enough events to account for all the causal truths (ii) causation *could* relate only facts, because all causal statements can be reduced to factive causal statements; and (iii) when we see how event- citing causal statements reduce to factive causal statements, we see that causation *really* relates the facts, not the events.

Mellor's paradigm causal statement links two Davidson-style event quantifications with "because":

(1) There was an F-event because there was a G-event

    for example, "there was a death of John because there was an avalanche". Assuming for simplicity that everything in the domain has name, (1) entails that there are singular terms "a" and "b" such that

(2) a = the F   and

(3) b = the G

    are true; which in turn entails that

(4) b caused a

    So, given that we new the names of every event there was, we could derive Davidson's transparent (4) from (1). However, the fact that a is an A and b is a B don't entail that

(5) There was an A-event because there was a B-event

    So, even if we knew all the types that a and b instantiated, we couldn't thereby conclude (5)or (1). So far, so good.

    Mellor then claims that what (4) does entail is that there are *some* types X and Y such that (6-8) are all true:

(6) b is the X

(7) a is the Y

(8) there was a Y because there was an X

Given that claim, we should agree with Mellor that this shows that "this shows how causation really relates the facts, not the events...the events merely inherit the causal relations of the facts they supply"(p.215). For even in the cases in which there is an event -citing causal claim corresponding to a fact-citing causal claim, we can reduce the former to the latter via some trivial identities. But to get from an event-citing causal claim to a fact-citing causal claim, we need independent knowledge of the property-level causal claim (8), which can't be gleaned from anything about the event s themselves. However, Mellor gives no good argument that I can find for the claim that for every event-citing causal claim, there must be some pair of properties for which 5-7 hold. If that's apriori, it's not transparently so; its denial appears to make sense.

This can be obscured by the fact that every event-citing causal claim of the form (4) can be expressed as

(9) a occurred because b occurred

But of course, that is not what Mellor intends, since the move from 4 to 9 is apriori (i.e. can be made simply on the basis of knowing 4) and has no tendency to imply that fact causation is always the only real causation.

In sum, then, Mellor's argument for the claim that causation relates only facts will ultimately be supported by the power of his analysis of causation. There is no strong case before the fact for insisting that there are no irreducible event-citing causal claims.

But even if Mellor had shown the latter, the former wouldn't follow. For Mellor has dealt only with the opposition between facts and events and there is at least one other apparently legitimate category of causal relata:

(a) Hitler caused the war

(b) The log caused the pile-up

(c) John caused the fight (or: caused a fight to break out)

(d) The goat caused the mess in your garden

All of these appear to be not only non-deviant, but commonplace forms of causal claim. (a-d) cite 'substances' - individual things that persist over time and that can be the subjects of predication but are not themselves predicable of anything - as causes.

Why have 'substantial' causal claims (henceforth 'SCCs') received so little attention ? one obvious reason is that most theories of causation don't apply to them; and as I have said before, this is fair enough. However, my immediate interest is in why such claims have not even been considered prima facie data for an account of causation. Note that I am not claiming that 'prima facie data' can just be read off surface structural form. The fact that "Mary's sake" is a surface syntactical unit in "I did it for Mary's sake" doesn't show that an ontological theory has to explain away the 'apparent reference' to a sake, for there is no such appearance. The statements (a - d) are more robust than this: they are worthy of being explained away. I will consider some points that have been made to this end.

(1) It is sometimes claimed that SCC's cannot be strictly and literally true, and/or must be elliptical. (see e.g. Menzies `89); the evidence being that, if pressed, a speaker will tend to give up her SCC, or fill it out, by introducing properties, states of affairs, or events. So, for example, if we respond to someone who asserts (b) or (c) with a query and a prompt - "well, it wasn't *just* John, in himself, was it ? Surely it was something he *did* ?" or "Really, you mean it was the *fact that* the log was left in the middle of the road..." - then they will tend to concur and re-express. However, this isn't a good argument for rejecting SCCs, for the fact that someone is willing to fill out a causal claim on request does not show that it was false, or even misleading, necessarily. It is just that if we are challenged in this way, it is natural to take it that the information we have proffered is incomplete or not sufficiently relevant in the context, that is, for the audience. That's a pragmatic matter. And, indeed, in the absence of background knowledge, SCCs are more apt to be of limited value than other forms; that is why they tend to be used only where the audience already has sufficient knowledge to see how x caused y. If I assert (b), one will immediately have a good idea of *how* the log caused the accident, i.e. by obstructing the road. But that doesn't mean that it wasn't *the log* that caused it.

To see that a filled-out causal claim need not eclipse the claim it replaces, consider the following case. I swerve across the road, causing an accident. When it is claimed that

(j) I caused the accident

what would be made of my response that it was *not* me, but rather

(k) my swerve/my suddenly swerving across the road etc caused the accident

The truth of (k) would not exclude me from culpability, because as we would say, it was me that swerved. (j) and (k) just aren't inconsistent. Similarly

(l) my moving the steering wheel caused the accident

(m) my decision to move my hands at *t* caused the accident.

and so on, are consistent (for different reasons) with (j). Our need to assign moral or legal responsibility is properly grounded in (j). And generally, whenever we use the locution " a caused X by/in virtue of F-ing", where 'a' refers to a substance and 'F' refers to a state, type of action, or property, "a caused X" is perfectly consistent with "a's being F caused X"

(2) It might be said that effects are always changes at times - either in a substance, or merely in a zone (e.g.) the coming about of a state of affairs or the loss or acquisition of a property - and that a thing cannot simply as such be the cause of a change. For given that a existed before and after the change, *without* having any effects at those times, there must have been some *difference* in the world at or before *t* to account for the change's occurring just when it did. Hence Fales (1990) (p.54):

" If as we typically believe, physical individuals persist through time and through change, it is senseless to speak of a particular simpliciter as a cause; for it is at one time involved in a causal interaction and at another time no longer involved in it. There would be no way to explain this fact if we could not refer to the differing properties had by that individual at various times during its existence. Nor could we explain why certain particulars enter into some causal relations and not others."

I fail to see how any of this shows that particulars (substances) cannot properly be said to be causes. Even given that every change can be caused only if there is a distinct

preceding or simultaneous change, why must "cause" apply exclusively (if at all) to that latter change, and not to what changes ? That would be to advocate something like Ducasse's theory, but not to argue for it. Explaining why a particular causes something at one time rather than another does not amount to claiming that it is not the particular, but something else, that 'really' does the causing.

Fales continues with "systematic grounds" for denying that causes are particulars: if they were, "it would not be possible to formulate causal laws, for the universality of laws is achieved by specifying not a list of particulars, but the properties they must satisfy to be the kind of things they are. And if we say that the particulars must be of a certain kind, the claim that causes and effects are particulars becomes indistinguishable from the one which takes causal relations to be individuals in virtue of certain of their properties. But the latter phrase can reasonably be understood as expressing what are event or states of affairs" (p.54-55)

This really doesn't add much to the foregoing. For (i) let us suppose that causation does require laws, and that laws relate properties. It doesn't follow that the causal relations that hold in virtue of laws relate the same kinds of thing s as the laws. And again (ii) to say that particulars must be of certain kinds or types to be causal relata is not at all say that the causal relata are really states of affairs, that is objects exemplifying properties.

(3)   Fales holds that causation is perceivable; and he thinks that this yields a phenomenological argument against taking particulars to be causes:
"When an object affects one in some way - say by striking one - it is never merely the object as such which one perceives as producing the effect. What is perceived is that certain of the object's qualities, and not others, are involved."

I agree with the second claim, but not the first. Just because I perceive that the motion and relative location of the ball are 'involved' does not show that I do not perceive that the ball simpliciter is pressing in to me.

So far, then, we have a good reason for, and no good case against, accepting particular things as causes, and so accepting the straightforward propriety of SCCs.

The justification for ignoring them in a story of causation must be simply the success of that theory for the cases it does apply to.

(III) PROBABILISTIC THEORIES OF CAUSATION

Let us now take up again the analysis of singular causation. We hit on two problems
for the reductionist theories considered in (I). First, there was the problem with
Menzies' "second platitude": that any given case of causation involves an *'intrinsic'*
relation that supervenes on the intrinsic properties of the relata and the intrinsic
relations between them. Both accounts revise our concept of causation, to a greater
or lesser extent, on precisely this point; and this seems to be reflected in the problems
with pre-emption, joint effects and causal direction. Mackie, as we saw, cannot
overcome these problems. Lewis, however, has a way of dealing with at least the
basic versions of these counterexamples, which appeals to the transitivity of causation
and his 'no-backtracking' principle (NBP). But this, as we shall see, faces serious
problems. Second, there was the general problem of indeterminism. Current best
opinion tells us that determinism is very likely not true of our world; but even if it *is*,
no philosophical theory of causation should presuppose that it is. As Lewis says, no
philosophical theory should have the consequence that if the world turns out to be
generally indeterministic, there is no causation. It is clear how this affects Mackie's
theory. Science, we are told, suggests that there are cases in which A causes B at t,
and yet everything that is the case at t -- whatever laws and particular facts hold at
t—is consistent with the *non-occurrence* of B at t. And common sense doesn't appear
to have any problem with this notion. Lewis's theory is also undermined. On the
CTFC, the status of a given sequence is causal will ultimately depend on there being
yes/no answers to questions of the form "if e had not occurred, would e2 have
occurred ?" But what if the only answer is "maybe, maybe not?" What we are being
asked to countenance under the hypothesis of objective indeterminism is that this
answer might reflect *not* an epistemological limitation—i.e., our inability to *tell*—but
objective facts about chance, that hold independently of inquiry. So, putting it

intuitively, it could be that e2 has some slight independent chance of occurring, regardless of e1, although e1 gives it a much *higher* chance of occurring. If e1 and e2 both occur, we want it to be coherent that e1 caused e2, even though some of the closest worlds in which e1 did not occur are worlds in which e2 did occur. And that's causation without counterfactual dependence.

There are various forms that a probabilistic theory of causation can take. The first thing to note is that a theory of causation can be probabilistic in the sense of allowing for non-determining causes, without actually analysing causation in terms of relations of probabilistic dependence (Salmon `84 and Menzies `96 present such approaches). In this chapter, we will be concerned only with those theories that do aim at an analysis in terms of probabilities.

There are then three main choice points. First, there is the kind of probability in question. We are dealing with *singular* causation: can this be understood solely in terms of probabilities attaching to types or properties, or do we need single case chances? Second, is the theory to use counterfactuals about probabilities, or traditional conditional probabilities? Third, is it required that causes "raise the chances of" their effects? This might seem to be essential to probabilistic theories, but Lewis denies it.

I will begin by presenting the most basic kind of probabilistic theory (the principles are taken from Salmon `80, `84; Suppes `84; Eells `91). Where F, G, and H are *types*:

- F is a *positive causal factor* for G if prob (G/F) > prob (G/~F)

- F is a *negative causal factor* for G if prob (G/F) < prob (G/~F)

- F is a *neutral factor* for G if prob (G/F) = prob (G/~F)

Then, the core condition on singular causation relating two events, e1 and e2 is:

(A) There are types F and G such that e1 occurs before e2, e1 exemplifies F, e2 exemplifies G, and F is a positive causal factor for G.

(A) defines a 'prima facie' cause (notice the need for the explicit temporal asymmetry requirement); but we need a further condition to exclude 'spurious causes,' which arise in cases of joint effects and epiphenomena. This is solved via:

(B) There is no pair H, e3 such that e3 is before e1, e3 exemplifies H and $prob(G/F \wedge H) = prob(G/{\sim}F \wedge H)$

This might be called Probabilistic neo-Humean theory; and it fails for much the same reason as Mackie's deterministic version did.

Recall our two marksmen Smith and Jones from ch.(I). Let us suppose that they are shooting at a target through a force field with the following properties. If a bullet hits it, there is a 50% chance that the field will allow it through, and a 50% chance that it will stop it dead. Within any 1 hour period, it can allow at most 1 bullet to pass; once a bullet has gone through it, the chance of a bullet passing through it drops to practically zero for the next hour. Smith fires (e1, at t1), in such a manner H that his bullet would be guaranteed to hit the target (prob ≈ 1), were it not for the force field; so the probability of the target being hit (G) is 0.5. Jones fires an instant later (e2, at t2) in such a way (F) that the bullet is on a direct course with the target. Suppose that Jones's bullet overtakes Smith's, hits the field first, and chances to get through, hitting the target; whereupon the field goes into prevention mode, stopping Smith's bullet in its tracks. Clearly, Jones shot, e2, caused the piercing of the target, e3, and Smith's (e1) did not. But the theory cannot account for this. e1, e2, and e3 exemplified types H, F, and G respectively, such that e2 counts as a 'spurious' cause (it's screened off by e1) of e3, and e1 counts as a genuine cause of e3. The crucial threat that we should note for probabilistic theories of causation here is that *raising of the effect's chances* is apparently *unnecessary* for causation here. On this construal of what it is for x to raise the chance of y, e2 fails to raise the chance of e3. And again we see the appeal of Menzies "2nd platitude." For the everyday epistemology of singular causation is much clearer than its philosophical analysis. In this case, we

have an ideal model of causation in the perceivable progress of the bullet, from the gun to the target; and, intuitively, nothing could impugn the causal status of Jones's shot unless it somehow affected that progress. In that sense, causation 'locally supervenes' on the spatiotemporal region between Jones's shot and the piercing of the target. But probabilistic relations clearly do not locally supervene in this fashion: we can affect whether Jones's shot raises the probability of e3 simply by hypothesising a 'spoiler,' e1, quite unconnected with that region. As we shall see, this thought is a central one in Lewis's and Menzies' theories.

The problem above was raised via a classic "pre-emption" case, familiar from deterministic theories of causation. But much of the discussion about probabilistic theories has focused on other kinds of case, of which there are innumerable variations on the theme. Here are two classics:

(A) The Squirrel (Eells and Sober `83; orig. Rosen `78)

I drive the golf ball (this is e1, of type F), resulting in an 80% chance of the ball's going into the hole (G); improbably, a squirrel emerges and kicks the ball (this is e2, of type H), deflecting it from its path Nevertheless, the ball travels from the deflection into the hole (e3).

(B) The Defoliant (Cartwright `83: Ch.2)

At t, a plant has an 80% chance of being alive at t3. I spray it with a killer defoliant at t2, lowering the chances of it's surviving to 20%. Nevertheless, the plant is alive at t3.

In case A, it can plausibly be claimed, the event that lowers the chance of the final outcome *causes* that outcome. The strength of the intuition is due to the fact that the epistemlogy of causation has little to do with probabilistic relationships: it is rarely the case that tacit judgements of probability increase are relevant to our making a causal judgement. In case A, we just *see* the model of a causal line traced by the ball, from one occurrence to another; that is why I find it hard to concur in the contrary results of the chance-raising probabilistic theory of causation.

Case (B) adds a puzzle for probabilistic theories. For, while it appears to exhibit no relevant difference in probabilistic structure to the other case, we intuitively *deny* that this is case of the spraying causing the survival.

There are endless ways of adding detail to the cases, and these ways of refining our causal judgements can provide bases for the chance-raising theorist to compromise on, or explain away, the apparent counterexamples. I will start by mentioning a few responses that don't do the trick.

(i) It might be said that the appearance of causation is genuine, but the appearance of chance-lowering is merely appearance: if we specify the putatively chance lowering event in sufficient detail, taking in i.e. the 'background context' of the exact angle and force of my drive and the exact physical details of the squirrel's deflection, we will see that the kick did in fact raise the probability of the final outcome (Salmon `80 calls this "the method of more precise specification"). But, as Salmon says, there is no reason to suppose that this will generally work. Although we cannot stipulate that these are cases of causation, surely under the assumption of genuine indeterminism, we can stipulate that they are case of physically irreducible low (and lower than before) probability.(Eells `91 p.284)

(ii) A related response points out that there will always be *some* type G′ instantiated by the outcome event, such that the putative chance-lowering event *did* raise the chance of G′(or: of the outcome event being G′).
But this isn't enough. Simply adding an effect that the theory *can* explain doesn't solve for the effect it *can't* explain. For example, even if we accepted, for some absurdly relational G′ that the squirrel's deflection caused the dropping of the ball to be (or 'caused it qua') G′ (say, 'a dropping two yards from a squirrel'), we would still have to account for the apparent fact that the kick caused the ball to drop.

(iii) The third response Salmon calls "the method of interpolated causal links": we locate events between the chance decreaser and the final outcome, such that chance raisings hold at each stage. But as the description of this solution shows, it

concedes that the kick need not raise the chances of (all) its effects. I will consider this idea further in the context of Lewis's (non-chance-raising) theory of causation.

Finally, the probabilist who demands chance-raising can simply brazen it out: these are not cases of causation, despite appearances. D.H. Mellor (1995) makes a qualified version of this claim, with an interesting supporting argument. To facilitate understanding of Mellor's case, we need to introduce Lewis's ('86) account of probabilistic causation - an approach quite different to that above.

The framework is easily understood in the light of the CFTC introduced in sec (I). The essential point is that we replace the core notion of causal dependence (CFD) with that of *probabilistic* counterfactual dependence (PCFD). When e and f are distinct actually occurring events such that the actual chance of **f** is x:

- **f** PCFDs on **e** iff (1) had **e** not occurred, the chance of **f** would've been some y such that (2) y is proportionally much lower than x.

As before, causation is defined as the ancestral:

- **e** causes **f** iff there is a sequence of events <**e1...en**>, each member of which PCFDs on its predecessor, such that **e1** PCFDs on **e** and **f** PCFDs on **en** (=df "**f** stepwise PCFDs on **e**").

The noteworthy features of the theory are, first, its appeal to single case, objective chances (as opposed to credences or frequencies) and, second, the use of counterfactuals with consequents about probabilities, rather than conditional probabilities. Lewis's probabilistic counterfactuals are evaluated as before, with the additional requirement that the chance x of **f** is to be read off the actual world, as it were, immediately after **e** occurs. This is to accommodate the time variability of chances.

The core of Mellor's probabilistic theory is substantially similar to Lewis's notion of PCFD. We recall first that the fundamental form of ordinary language causal statement is factive, in Mellor's view, and at this level of analysis, the relata of causation are facts. Every true causal statement can be represented as "P because Q",

where P and Q are whole sentences, and states-of-affairs are their ontological correlates. *Facts* are that subset of states of affairs that are correspond to *true* sentences. Thus it is to states-of-affairs that probabilities - single case, objective chances - attach.

The next complication is that Mellor relativises the chance of a state of affairs to other states of affairs: the chance of P is a property of some state of affairs Q. To see this, consider John, now at a party, who is planning to drive home later. What is the chance of John having a road accident tonight? It depends. When he arrived at the party, firmly intending to remain sober, it was low, say 0.05. But as John ploughs his way through the cocktails, it rises sharply. If he collapses unconscious, it will then drop again. We might capture this by representing "the chance that John has an accident tonight" as "ch$t$(P)", which takes different values according to the variable t. Instead, Mellor uses "ch$f$(P)", where "f" is a placeholder for different states of affairs. So, for example, if Q = "John has drunk four vodka martinis", "ch$Q$(P)" might be something like 0.8; if R = "John is catatonic", ch$R$(P) will be near enough 0. The important thing for Mellor is that "ch$X$(Y)" is never the same thing as "ch(Y/X)", even though they may take the same value. "ch$X$(Y)" (that chance of Y that is a property of the state of affairs X) is to be read not as the conditional probability, but in terms of a Lewis probability counterfactual "the chance that Y has in the closest world in which X is true", or "[ιp] [X$\Rightarrow$ ch(Y) = p]". This means that whenever X actually obtains, "ch$X$(P)" is just the actual chance of P.

The essence of "P because Q", then, for facts P and Q, is:

[M] ch$Q$(P) > ch~Q(P)

i.e., the actual chance of P is greater than the chance P would've had if Q were not true, or again, the actual chance P is greater than the chance P has in the closest ~Q-world. Thus causation is, essentially, Lewisian PCFD: Mellor does *not* adopt the weaker Lewis notion of "stepwise PCFD". Now Mellor holds to the "truthmaker principle": every (contingent) truth requires a substantial ontological ground, something in the world that *entails* its truth (he calls these 'facta', to distinguish them from the 'thin' fact-entities that are analytically posterior to true sentences.) When we

find "what makes statements of the form[[M]] true" - and ipso facto, what makes statement as of the form "P because Q" true - we will have our hands on the "real embodiment of causation" (pp.168-9). This embodiment turns out to be "nomic facta", local instantiations of laws of nature.

To return to the problem for chance-raising requirement on causation. We have a counterexample to Mellor's theory (if and) only if :

(i) "P because Q" is true   and

(ii) "ch$Q$(P) > ch~Q(P)" is false

Mellor sticks to his guns and denies (i) holds for cases such as the Squirrel. Admittedly, this doesn't *look* plausible. Mellor's chosen case is an ancestor of that presented above: the golfer pulls his drive, resulting in the ball hitting a tree and from there ricocheting into the hole. The objection is that the mishit lowered the chance of the effect that it caused. recalling Mellor's distinction between particular and factive causes (see II), he can very plausibly avoid this problem. Where P = "John sinks the ball" and Q = " John hits the ball", (ii) is true, thus (i) is true; and so, if f = Johns shot and g = the sinking of the ball, f caused g is true. Further, since f = John's mishit, "John's mis-hit caused g" is true. The only thing Mellor must deny is "the ball sank because John mishit it", which is plausibly false. However, the squirrel case isn't so convenient. In Mellor's story, the chance-lowering bogus cause was the state of affairs of an event having a certain property: the drive's being pulled, or John's mishitting his drive. Since that property was *not* one in virtue of which the ball went into the hole (putting it intuitively), the causal claim which says it *was* comes out false.

In the squirrel case, we've got a particular cause (the kick), which is no problem for Mellor (I expressed doubt in the last chapter about Mellor's claim that events are always merely derivative, and not genuine causes. But I will take it for granted here.) What's the factive cause? Apparently just: the state of affairs of the squirrel's kicking the ball, or the fact that the squirrel kicked the ball. It is this which Mellor must deny causal status to. But, the property of being kicked by the squirrel does appear to be

one in virtue of which the ball goes into the hole; consequently "the ball went into the hole because the squirrel kicked it" is far from obviously false. Still, it's not obviously absurd to deny it, and Mellor does: the ball goes into the hole despite, but not because, the squirrel kicks it. (Notice that Mellor would have to concede that if the ball went into the hole after being kicked while just lying around stationary, it *would* be a case of causation).

What's interesting is that Mellor (`95, ch 6; `91) has an argument to show that our intuitions to the contrary *must* be misplaced. The idea is this. There are certain functions ('connotations') of causation so fundamental to it that nothing which failed to support them could be worthy of the name; but any pair P,Q will support them only if they satisfy [M]. That is: a relation is causation only if it supports the connotations, and a relation supports the connotations only if it involves raising of chances. Therefore our intuitions in the squirrel case must be leading us astray.

The three connotations are (1) the *explanatory* connotation (2) the *means/end* connotation and (3) the *evidential* connotation. Let us examine these in order.

(1) The Explanatory Connotation

That causes explain their effects is indisputable (see the discussion of Davidson in (II)). How does this require principle [M]? Mellor's argument (pp.73-77) is:

(a) Explanation works by (has as a principal function) closing the gap between what we now to be the case as a mere matter of fact and what we know to be necessarily so.

(b) Chances measure those degrees of possibility that are objective properties of facts about the world: the greater (lesser) the ch(P), the closer to (further from) being necessary P is.   Therefore:

(c) X can explain Y only if X raises the chance of Y - the more the chance is raised, the better the explanation.

The contentious premise is (a). I think that it is false, and derives its plausibility from the fact that *showing why something was more or less necessary*, which is rarely (if ever) the primary aim of explanation, is a natural accompaniment of something that *is,*

namely *showing how such-and-such came about*. We understand why P is the case by seeing how it came to be the case that P. Consider a few examples:

- Why did John hit Jim? Well, Jim had been ordering him about all day; and they argued about the washing-up last week; and they'd both had a drink...etc. Of course, I'll say less or more, in better or worse order, depending on what my audience knows about the situation /about people in general, but the point should be clear. The information I give will show why the event was more likely to happen *only because* it appeals to antecedently understood principles - causal and psychological - concerning how one thing tends to lead to another, that are instantiated here.

- Why is the window broken? Because the football hit it. Again, there is a story (perhaps partly to be told by the explainer, partly to be reconstructed by the audience) that lets us see *how* it came about - that appeals to a means by which the window was broken.

- Why is the ball in the hole? Here it seems even more evident the explanation will be of how the ball came to be where it is. Suppose I watch the golfing situation. I see how the ball gets from the squirrel's foot into the hole. So I have an explanation of how the ball got into the hole: because the squirrel kicked it in. Of course, "because the squirrel kicked it" might not make it pellucid to you. But that doesn't stop it being a perfectly good explanation. No-one would say that citing the cause alone necessarily makes for the *best* explanation, since quality of explanation is governed by pragmatic factors. In this case, the explanans-event cited didn't raise the chance of the outcome, but it let us see how it came about, so it explains it.

One might respond that this way of arguing begs the question against Mellor, since in these three cases, "showing how x came about", "means by which", "mechanism" etc. just amount to the claim that we explain by citing some of an event's causes; and if Mellor is right, causes must raise the chances of their effects. But:

(i) even if that's all there is to this response, that doesn't make it question-begging. It would merely mean that it is not likely to advance the dialectic - and it was Mellor's

responsibility to do that. The *point* of arguing from "connotations" is to give grounds relatively *independent of* any particular theory of causation for thinking that causes must raise the chances of their effects. If one holds (as Lewis does) that all singular explanation is explaining via causes, and one disagrees (as Lewis does) with Mellor about causation, the explanantory connotation won't resolve anything. (see Beebee `98)

(ii) I don't think that the notions appealed to against Mellor simply reduce to causation. To see this, consider the idea of explaining why x has a certain supervenient property, such as liquidity, in terms of its microstructure. On a Lewis/Jackson model (see e.g. Jackson `94; Chalmers `96), we would first give a conceptual analysis, C, of liquidity in terms of dispositions to behave in certain ways; and then show that a certain molecular structural basis entails that that the macroscopic body will satisfy C. In this way, the basis *determines* the presence of the supervenient property, but not *causally*. (Similarly for the philosopher's favourite explanation: of "x has temperature y" in terms of "x's molecules have mean kinetic energy K.") We have a conceptual handle on determining something to be the case that is broader than the notion of causing it. And this handle isn't a grasp of primitive chance-raising; but rather the grasp of a mechanism.

I conclude that the explanatory connotation at best does nothing to advance Mellor's case, and plausibly sits very well with the chance-lowering counterexamples.

(2) The Means/End connotation

"To call something a cause that provides no way of bringing about its effects seems to me an obvious contradiction in terms"(p.88). Mellor clearly takes this connotation to provide his master argument for the CRR. The strategy is as follows.

(P1) Causation is "what gives ends means"("the fact is undeniable") p.79

(P2) We can give an account of the concept of *a means to an end*, which

(i) doesn't presuppose causal notions and (ii) entails that means must raise the chances of their ends.

Conclusion: causes must raise the chances of their effects.

P2 is filled out in the framework of an objectivised, non-causal decision theory, in which 'subjective valuations'(i.e. the 'desire' parameter) and credences (the 'belief' parameter) are replaced by the notions of objective utilities and objective chances. So, with respect to some given end, G, of mine and an action-type, V-ing, we have the objective utilities of my V-ing and getting G, $u(V \wedge G)$, and of my V-ing and not getting G, $u(V \wedge {\sim}G)$; and the objective chances of my getting G if I V, $chv(G)$, and of my not getting G if I V, $chv({\sim}G)$. We then define the *mean* (as in average) *utility* of my V-ing as

[M1] $mu(V) = chv(G) \times u(V \wedge G) + chv({\sim}G) \times u(V \wedge {\sim}G)$

and the mean utility of my not V-ing as

[M2] $mu({\sim}V) = ch{\sim}v(G) \times u({\sim}V \wedge G) + ch{\sim}v({\sim}G) \times u({\sim}V \wedge {\sim}G)$

Now:

(a) the *mean utility principle* (MUP) - an objectivised version of the priniple of maximising subective expected utility - prescribes that I should do V iff $mu(V) > mu({\sim}V)$, that is iff the overall utility for me of V-ing exceeds the overall utility for me of not V-ing.

(b) we assume that Ving is a 'pure' means, that is, has no intrinsic value or disvalue of its own: $u(V \wedge G) = u({\sim}V \wedge G)$ and $u({\sim}V \wedge {\sim}G) = u(V \wedge {\sim}G)$.

Then

(c) Since, ex hypothesi, $u(G) > u({\sim}G)$, it is easy to show that: [the MUP will prescribe thta I should do V] iff $[mu(V) > mu({\sim}V)]$ iff $[chv(G) > ch{\sim}v(G)]$.

(d) Finally, assume that what it is to be for arbitrary φ to be a means to a given end E *is* for the MUP to prescribe that I should do φ.

Given (P1), we have the conclusion: Y because X $\rightarrow$ $chx(Y) > ch{\sim}x(Y)$. Causation must satisfy the CR-condition [M].

What's wrong with the argument?

First, remember that what Mellor is trying to do with the argument from connotations is get some independent conceptual purchase on the chance-raising requirement; P1 counts as undeniable just to the extent that the explication of *means* is genuinely a elucidation of the intuitive notion. But (d) doesn't look at all like a plusible explication from this point of view. Let me explain.

(i) Mellor cannot, without, sacrificing his argument, insist that there can be only one means to a given end. For it is not with respect to *that* notion of means that the "undeniable" connotation - "causation is what gives ends means" - is undeniable. Further, one means to a given end can be better or worse than another. This I take to be obvious. One can get oneself a broken window (G) *either* by throwing a ball at it (V1) *or* by serving the ball at it with a racket (V2), but you're more likely to succeed by the latter method - that is, it's the better means to that end.

Now at first sight, Mellor should have no problem with this. Since either of $V1$ or $V2$ will raise the chances of (G), both can count as means to that end, according to M1, M2, and (d). However, suppose

(1) the circumstances are such that both means are available, but either one will operate iff the other does not. For example, we might have a single ball, which I want to throw and you want to serve. Since

(2) $\mathrm{ch}v1(G) < \mathrm{ch}v2(G)$

it follows that

(3) $\mathrm{mu}(V1) < \mathrm{mu}(V2)$

and so, that, in these circumstances

(4) $\mathrm{mu}(V1) < \mathrm{mu}(\sim V1)$

Therefore, the MUP will not prescribe that I do V1, therefore by (d):

(5) $V1$ is *not* a means to G - contrary to (1).

Now this would be question-begging as a reductio of Mellor's account of 'means'; for Mellor can simply deny the premise (1) that this is a situation in which we

have two available means. Rather, *V1* fails to be a means in this context precisely because it does not raise the chances of the effect. I find this deeply implausible. As with explanation, the fact is that chance-raising isn't conceptually basic in our understanding of what it is to be a means to an end. Providing a mechanism, or a way of bringing about, is. Again, Menzies' platitude applies. Kicking a ball into a hole is a basic mechanism for geting it into the hole. The mechanism can't be affected by what goes on at other spatiotemporal locations, unless such goings-on affect the intrinsic processes at the location where the mechanism operates. But we *can* change its status as a *chance-raiser* without interfering with those intrinsic processes at all. Our notion of a means tracks the mechanism, not the chance-raising. So, we can accept the Means/End connotation without accepting that causation must satisfy [M]. At any rate, it is clear that the question of whether means must raise the chances of ends is really no clearer than the question of whether causes must increase the chances of their effects, and so Mellor's appeal to the means-end connotation does nothing to advance the dialectic.(see Beebee `98 for an earlier statement of this objection.)

(3) The Evidential Connotation.

Again, we start with an evident platitude: cause and effects are *evidence for* each other.(cf. Hume's view that causation is the ground of all our beliefs about the world-beyond-immediate experience.) How could causes be evidence for there effects without raising their chances, without, that is, satisfying [M]? Clearly, if X makes no difference whatsoever to the likelihod of Y, or X make sy *less* likely, then one couldn't take X as evidence for Y. So evidence requires chance-raising; therefore, so does causation.

However, this is the weakest connotation. For anyone (like Mellor) who takes an objective view of causation, in the sense of holding that the causal facts obtain regardless of anyone knowing or believing that they do, should be sceptical of tying causation to such an obviously context-sensitive epistemic notion as evidence. Considere the following propositions:

(i) X took drug A at t1   **(ii)** Drug A has a 90% chance of inducing a fatal heart-attack.   **(iii)** X took drug B at t2   **(iv)** Drug B has a 1% chance of inducing a fatal brain haemorrage **(v)** Drug B has a 99% chance of neutralising drug A **(vi)** X died at t3 of a brain haemorrage.

Suppose you know only (iii) and (iv). Then you will take (iii) as evidence that X died. Whereas, if you know only (i) and (ii), and (v), you will take (i) as evidence that X died. If you then discover (iii), (iii) is evidence that X does not die. But given that (iii) causally explains (vi), we can't change the causal status of (iii) simply by adding to someone's information. Again Menzies point applies. There's an intrinsic physico-chemical mechanism by which drug B induces brain haemorrage, which operates in Jones between t2 and t3. The truth of (i) and (ii) can't affect that, and so can't affect whether (iii) causally explains (i). To put it another way, causation supervenes much more locally, at least in some cases, than evidence. Evidence tends to be governed by observed frequency relations amongst types - not single case chances. But clearly, empirical frequencies don't supervene on the immediate location. (Again, this kind of objection was first put - rather more clearly - by Beebee `98).

In sum, then, Mellor's arguments for the chance raising condition from the connotations of causation are not at all suasive. The condition has implausible consequences that we have no reason to accept. So we should reject the condition. I think we can draw a more specific conclusion from the examples considered. I have appealed to cases which illustrate Menzies locality platitude - that is, cases in which, intuitively, causation supervenes on the intrinsic physical facts concerning a certain spatio-temporal region. Those intuitions are matched by commonsense causal judgements. Take the squirrel case again: we see the ball travel from the impact with the squirrel into the hole. Now clearly, facts about chance-raising need not supervene in this way in these cases, but may depend on, for example, what obtained beforehand. Therefore, in any such case, causation can come apart from chance raising. Recall the pre-emption case. The fact that Jones fired does not satisfy [M] with regard to the fact that the target is pierced - simply because Smith fired first. But insofar as *firing a bullet into* something is a way of causing that thing to be pierced, it

cannot lose that status simply because someone else fires first. So it is very hard to see how any theory that requires causes to raise the chances of their effects can succeed.

This is why Lewis's theory, to which I now turn, does not require it. For Lewis, stepwise PCFD is sufficient for causation. Hence he has no problem with the squirrel case, since we can identify a sequence of events, between the event that lowers the chance of the final outcome and that outcome, such that each event PCFD's on its immediate predecessor. (Lewis isn't too conceerned with everyday criteria for what counts as an event, as we saw in ch.(II): the golf ball's moving in direction D with velocity V at time t will do fine for him.)

The problem for him is the probabilistic analogue of the pre-emption case evaded in ch (I). Under deterministic assumptions, the basic idea was that there are two events, c1 and c2, and an effect e3 such that c1 causes e3, but had c1 *not* occurred c2 *would've* caused e3. The problem for Lewis was that, if every *ci* that is causally intermediate between c1 and e3 is such that if it had not occurred, c1 would not have occurred, then there is no way of setting up a chain of CFD from c1 to e3. For any ci, if it had not occurred, then c1 would not have occurred, and c2 would've caused e3: e3 does not stepwise CF-depend on c1.

Correspondingly, the prima facie problem presented without the deterministic assumption is that no stepwise PCFD-chain can be constructed from the pre-empting cause to the effect. The Smith and Jones case above illustrates probabilistic early pre-emption: the 'main line' from the pre-empting cause (Jones shot, c) interferes with the alternate line from the pre-empted cause (Smith's shot , d), *before* the final effect. If events can PCFD on their temporal successors, then there is no way of constructing a PCFD-chain from c to e. Suppose that for every ci causally between c and e, we can reason counterfactually that if ci had not occurred, then *ci-1* would not have occurred, then *ci-2* would not have occurred...etc, all the way back to "c would not have occurred". Then there is *no* cn such that: if it had not occurred, *the chance of e would have been lower*. For d would always be there to confer the same chance on e.

Lewis's solution, as we saw, involved the 'no backtracking' principle NBP for evaluating counterfactuals. Take some c*n* that is the event consisting of Jones's

bullet's motion an instant before it hits the target. By the NBP, the closest world in which cn does not occur is a 'minor miracle'-world, w, in which every stage of the bullet's progress between $c1$ and cn *did occur.* So, in w, Smith's bullet has already been stopped at the time at which cn fails to occur, and the chance of the target's being pierced is practically zero. Which is just to say that at the actual world, e PCFDs on cn; which allows us to reconstruct stepwise-PCFD between c and e.

Now things get much stickier. First up is the problem of "late" pre-emption, mentioned in ch(I). The idea is that there be 'blocking' pre-emption, *but with no event prior to the occurrence of the final effect* that blocks the pre-empted line. In these cases, the NBP is to no avail. For the success of the NBP crucially depends on their being some time (and at least one event) at which (i) the pre-empted line has been blocked and (ii) the final effect has not occurred. It is only events from that time-slice on which the final effect will PCF-depend. So, for example, simply alter the Smith and Jones case so that the force field goes into prevention mode *not* when *it* has been pierced, but only when the target itself has been pierced. There is now no cn causally between c and e such that if cn had not occurred, the d-line would already have been blocked; hence there is no cn such that if it had not occurred, the chance of e would have been lower, and we cannot reconstruct stepwise PCFD on from c to e.

Second, there is a serious question about the validity of NBP- and so about the viability of the solution to the early pre-emption problem, which we have thus far accepted - without the deterministic assumption. Remember that the NBP is justified by Lewis's weighting system for similarity of worlds, which tells us to preserve substantial matches of particular fact at the price of small violations of law. Given that everything is deterministically caused, this generally rules out "backtracking", since the only way to absolutely avoid violating any laws in making a counterfactual supposition would be to alter the entire history of the world prior to that event. Now our assumption of indeterminism presumably allows us for a "mixed world": a world involving both deterministic and indeterministic causal links. Imagine a pre-emption set-up meeting the following specifications:

(i) There are three events - c1, c2, and c3 - on the main causal line to e.

(ii) c2 is the 'blocking' event that cuts off the alternate line leading from d to e.

(iii) The causal link between c2 and c3 is *deterministic,* but the connections between c1 and c2, and between c3 and e are both *chancy.*

Suppose that c3 had not occurred. What obtains at the closest world to this ~c3-world? According to NBP, this is a miracle world, at which c2 somehow fails to produce c3, but was nevertheless there to inhibit the d-line. So, if we *flout* NBP, we avoid 'miracles' - and respect the 'iron' connection that actually links c2 and c3. What about particular facts? Well, it seems that we don't need to change any *more* of these at all. Ex hypothesi, the link between c1 and c2 is actually merely probabilistic - so, in getting rid of c2, we can consistently suppose that c1 still occurs, *without positing any miracles.* So, the choice is, for counterfactual reasoning about c3:

(a) we can allow 'backtracking', thus avoiding any law violations, at the price of one particular fact (that c2 occurred).

<div align="center">or</div>

(b) We can stick with NBP, saving all the particular facts, at the price of one law violation.

To follow NBP would amount to claiming that a single particular fact counts for more, in the similarity stakes, than the violation of a deterministic law; which doesn't seem defensible, even by Lewis's own principles. But if the NBP fails, then so does Lewis's solution to the "early" pre-emption problem.

The situation, then, is this. the CFTC faced a problem about early pre-emption, that was avoided via a contentious, but not implausible principle NBP. However, under indeterminism - or at least, a 'mixed' world - NBP lacks its original motivation[6]. So, either the Early pre-emption problem remains unsolved, or Lewis concedes that he is taking a gamble on absolute indeterminism - no mixed worlds - which we know he does not want to do. Worse, even if he takes the latter option, their remains the Late pre-emption problem, which, although it appears to falsify the CFToC for the very same reasons as the Early pre-emption problem, cannot possibly be solved by the NBP. So it appears that Lewis is going to have to come up with a

---

[6] Dr. David Papineau has since pointed out to me that my 'mixed world' example in fact fails undermine the NB principle, which is actually grounded in the need to respect the 'asymmetry of over determination': see Lewis's 1979.

different kind of solution to the Late pre-emption problem, further weakening the CFTC. We might suspect that if the theory has to give different solutions to what is essentially a single kind of problem, it cannot be right.

Here the story takes a radical twist. Lewis (`86b: pp.205 - 207) proposes a solution to the Late pre-emption problem that, if it works at all, should work for *all* the pre-emption cases. In effect, Lewis bites the bullet on Menzies' platitude. What prevents stepwise PCFD from holding in the 'right' places in all the pre-emption cases is always something quite 'extraneous to' whatever is going on - the intrinsic processes'- between c_3 and e; and such extraneous factors, he notes, cannot interfere with the causal facts. Consider the last mentioned pre-emption case. Lewis invites us to consider *other* spatio-temporal zones - actual or merely possible, but with all the laws held fixed - that are *qualitatively indistinguishable from* the smallest zone containing c, e and the processes that go on between them, *but which are not embedded in the contexts of 'rival' processes*. Here, PCFD *does* hold between the events corresponding to c_3 and e (Lewis actually presents this solution in the context of plain old CFD. But presumably it is of value only if it works for the probabilistic account.) But any one of the zones exhibits a causal connection iff they *all* do. So we have:

(1) f *quasi-probabilistically counterfactually depends* (QPCFD's) on e iff there is a possible zone z in which events e' and f' are linked in accordance with the standard CFTC-account; and there are processes connecting e and f, the intrinsic character of which is the same as that of the processes connecting e' and f'

(2) e and f are linked by a QPCFD-chain iff there is a sequence of events <e1...en>, each member of which QPCFDs on its predecessor, such that: e1 QPCFDs on e, and f QPCFDs on e*n*

(3) e causes f iff there is a QCFD-chain linking e and f.

The first thing to note is that this response cuts quite through the tangle of technicalities woven in the foregoing. Although Lewis does not present it in this way, it gives a uniform method of dealing with the pre-emption cases. For if it is legitimate, why *bother* with the NBP? After all, it would certainly be easier on Lewis to avoid staking his theory on a disputed principle.

Do we have reason to think that this solution is watertight? It is hard to seee how there could be an LP or EP counterexample to this theory, since it effectively amounts to principle way of *stipulating away* any factors that might interfere with CFD. My real objection to this account is that it doesn't *save* the *CFTC*; it precisely shows that PCF-dependence merely provides excellent but defeasible *evidence for* causation. (Which is what the vulgar view held all along: causal dependence *explains* counterfactual dependence)."Quasi-counterfactual dependence" clearly isn't a *species of* counterfactual dependence; it's just a way of picking out causal dependencies that don't support counterfactuals.

Menzies (1996), as we shall see in the next chapter, takes Lewis's idea to its logical conclusion: locate causation in those intrinsic processes that Lewis appeals to, not in the counterfactuals that they sometimes fail to support. Whatever merits there are to Lewis's theory will be taken over by Menzies' theory. I conclude that there is no reason to think that any CFTC can adequately account for probabilistic causation.

The last probabilistic theory of causation that I will consider is presented in E.Eells `91, and is quite different from any we have considered thus far. On the one hand, Eells agrees with our evaluation of the basic neo-Humean probabilistic theory: probabilistic correlations between types just don't suffice to fix the token level (singular) causal facts, so no type-level theory of causation will apply directly to the token level. On the other hand, Eells makes no use of the single case chances or counterfactuals: the theory doesn't use any probabilistic tools other than conditional probabilities relating types.[7]

Eells doesn't think that we can use comparisons of conditional probabilities of the form prob(X/Y) > prob(X/~Y) in explicating singular causation. In the first place, we can't evaluate prob(X/~Y) if Y is a particular fact: if John actually does study logic, the probability that he does not study logic is zero, and prob(X/~Y) (i.e., prob(X&~Y)/prob(~Y)) goes undefined. In the second place, we have the cases considered above (the squirrel etc.) in which the F-event causes a G-event even

---

[7] To be more precise, Eells explicates probability as frequency in hypothetical populations of a type exemplified by our target token population. Since his theory of singular causation can be explained without appeal to this idea, I omit the details here.

though the *type* F is a *negative* causal factor for the type G. The key to singular causal relations between e1 and e2, on Eells' view, is the notion of a *probability trajectory*: how the probability of e2 "actually evolves around the time of the earlier event e1 and between the times of e1 and e2."

Before expanding on this, we should note that Eells' view of singular causal relata isn't quite like any considered in ch(II). On the one hand, Eells' 'events' are quasi-Davidsonian, "concrete" entities in that a given event is fully specified by and only by a spatio-temporal location plus all the properties exemplified at that location. However, since we can evaluate the probability of an event only if we can discern the relative frequency of some type F in hypothetical populations of some kind H, we must take events as exemplifying kinds. This doesn't involve us in saying e causes e2 "qua such and such" or "under description D", because it is not events that are fundamental causal relata, but exemplifyings-of-types by events. Eells' fundamental causal locution must be "x's exemplifying X causes y's exemplifying Y", or "y exemplifies Y because of x's exemplifying X" - *not* Davidson's "x caused y" or "y because of x".

Eells defines four kinds of causal significance relations, in terms of four basic probability-trajectory-schemata. The two that are relevant to our discussion are[8]:

(1) y is G *because of* x's being F if: (a) the probability of G changes at the time of x(=tx) (b) immediately after x, the probability of G is high, and higher than before tx (c) the probability of G remains high until the time of y (=ty).

(2) y is G *despite/in spite of* x's being F if: (a) the probability of G changes at the tx (b) after x the probability of G is low, and lower than before tx.....the probability of G may o may not recover by ty.

We can make this more intuitive by applying it to the squirrel case. Notice that the definitions 1 and 2 make no reference to the prob(G) *at* the time of x, but only to what happens before and after; so the statistical relation between being kicked by squirrels and going into the hole doesn't matter. What *does* matter is that, since in the case as described, the ball comes off the squirrel's foot on course for the hole, the

---

[8] The other two are 'independently of' and 'autonomously of'

probability of its holing went up immediately after the e vent that was a kicking, and stayed up until the outcome. So Eells can account for the positive causation here. Note also that if we instead suppose (as some versions of the story do) that the squirrel kicked the ball *away from* the hole, the ball only reaching the hole by way of a further deflection from a tree, Eells' theory gets the plausible result that the ball went in to the hole because it hit the tree, but *despite* the kicking.

The theory is further confirmed by the way it discriminates between Cartwright's defoliant case and the squirrel case. In the framework of the neo-Humean probabilistic theory, recall, the two case were structurally identical: in each case, we have a development, from the instantiation of a factor F negatively relevant to a factor G, to the instantiation of G. In one case, the development intuitively supports a positive causal relation, whereas in the other, the development intuitively has the negative causal significance of 'despite'. The probability trajectory method gets the right result: the survival of the plant (y's being G) was despite the spraying of the plant (x's being F)[9].

In our pre-emption story, we have Smith's shot (z's being H), Jones shot (x's being F), and the piercing of the target (y's being G). To assess the causal significance of Jones's shot, we check how the prob(G) evolves around the time t_x of Jones's shot, keeping in mind conditions (i-iii). To our original case, we add the stipulation that: both Smith and Jones fired on a snap decision, so that in each case, it was very unlikely that the shot would occur until it did.

How does prob(G) look just prior to Jones's shot? Well, Smith has already fired a well-aimed shot, so that the prob(G) ≈ the probability of a bullet getting past the force field, say 0.75. How does prob(G) look just *after* t_x? It doesn't change: since the field will allow at most one bullet through, prob(G) is still 0.75. So by the trajectory schemata, Jones's shot comes out classified as causally independent of the

---

[9] Eells goes on to add three further conditions specifying what must be 'held fixed' in evaluating for causal significance. I omit them, first, because they do not affect my case against Eells and second, because all three use explicit causal notions and thus seem to make his account hopelessly circular.

target's being hit, which we know is false. Eells' account isn't necessary for causation.

The problem is that Eells' theory can fail in any situation in which there is some *upper limit* on the value that the probability of the outcome can take before the fact, so to speak. So, as Eells concedes throughout, his theory is viable only on the assumption that determinism is false; and that means *any degree* of determinism. The point is simply seen by considering the idea of a "mixed world" that we deployed against Lewis's 'no backtracking' principle. Suppose e1 deterministically causes e2, which chancily causes e3 (the exemplification of G); that is surely a non-question begging characterisation of a possible situation. We just cannot account for the link between e2 and e3 by Eells' method of evolution. Given that e2 is guaranteed to happen before it does, whatever probability e3 has after e2, it has prior to e2, so the trajectory of prob(G) won't change around t2 and e2 will appear to be causally independent e3.

Now since Eells' theory of "because" crucially relies on the notion a rise in probability, the problem that the deterministic link would make for it is simply that it precludes the relevant prob(G) from rising. And this is fatal, because it opens Eells' theory to counterexamples even under his assumption of global indeterminism, as the Smith and Jones case shows. Prob(G) can have an upper bound set by mere (perhaps nomological) matter of empirical fact; we can suppose it's just an a physically irreducible probabilistic fact that the probability of a bullet getting past the force field is 0.75. Once Smith's shot raises the prob(G) to that limit, there's no room for Jones's shot to raise it any more, and so no room for Jones shot to cause the instantiation of G. Yet it does cause the G.

From a still more abstract point of view, Eells' theory fails in the same way that Mackie's, Mellor's and Suppes's did. We know that Jones's shot causes the final effect, because it's Jones's bullet that travels into the target. This is on case in which the causal facts supervene on the physical facts at a (fairly clearly delineable) spatio-temporal location. Short of a miracle, nothing could affect the causal status of

Jones's shot without affecting the intrinsic goings-on at that location. But whether the probability of the final effect changes around the time of Jones's shot *can* depend on factors that do not affect those goings-on (e.g. Smith's shot) So probability change can't be essential to causation. In the next section, I finally turn to some theories that take this point to heart.

(IV) CAUSATION AND PROCESSES[10]

Let us briefly establish where we stand at this point. I have surveyed a variety of theories of causation and found them decisively unsatisfactory. The first general disadvantage is that they all come with a certain prejudice about the proper relata of causation: some restriction on the kinds of thing that can be causes and effects that is required by the theory, but inadequately defended. However I decided that is hard to refute a theory on this basis (depending on the severity of the restriction it imposes) since a little re-negotiating of what we would ordinarily say -- for example, that a person or other object caused something -- is a price that one might pay for a theory that is otherwise successful. But therein lies the second problem: these accounts simply don't always get the right results, even on their own territory. Although it is undesirable to argue purely by counterexamples, I have tried to make my objections as general as possible, by showing, in each case, that given the form of the problematic example and the materials available to the theory in question, we have good reason to think that the theory *cannot* solve it. In particular, I have leant on the idea, due to Menzies, that -- in some cases at least -- we ordinarily assume that a causal relation between events (say) involves some connection between them which is determined by the intrinsic natures of those events and of the local spatio-temporal region between and including them. This intuition is given some kind of concrete form in the case of the squirrel and the case of the rival marksmen. In the former case, we *see* the ball travel from the squirrel's foot into the hole; this suffices for us to judge it to be a case of causation. But Mellor's theory ignores the obvious, or rather registers it only indirectly: that is, as just another aspect of the situation which may explain the truth or falsity of the chance-raising counterfactuals that are the real essence of the causal claim. As a result, the theory is apt to give the wrong result. For although my kicking a ball into the hole is as clear a case of causation as one could want, whether my kicking the ball (i.e., a fact) raised the chances of its going into the hole depends on factors quite extraneous to its causal status -- for example, whether the ball was

---

[10] Since completing this thesis, I have been told that a more promising (and scientifically informed) process theory than those covered here has been developed by Phil Dowe.

stationary or moving prior to my kicking it. Lewis faced similar problems with pre-emption that led him, as we saw in the last section, to concede -- effectively -- that causation is not *simply* a matter of counterfactuals, but also involves something else that typically, but not universally, *supports* such counterfactuals.

In this chapter, I deal with a family of approaches that aim to explain causation in terms of Lewis's 'something else.' This might be called the "process" theory of causation, although, as we shall see, "process" is here but a place-holder for a notion that needs a theoretical characterisation. The common factor is the positing of some relation between causes and effects which is (1) *intrinsic*, in a way that probabilistic and counterfactual relations are *not*; and (2) can be used in an illuminating, non-circular account of what causation is. However, this kind of theory involves a different take on analysis than has so far been assumed. We can introduce this, and give some structure to the plot of this discussion, by mapping theories of causation against the recent history of the philosophy of mind.

The theories examined so far are engaged in a certain kind of conceptual analysis: they aim to give reductive truth-conditions for statements like "A causes B". Given a satisfactory answer, there is taken to be no further empirical question about 'the nature of the causal relation.' We can compare this approach with logical behaviourism, which saw the whole truth about the mind in the reduction of psychological statements to non-psychological conditionals relating behaviours to circumstances, and denied that there were any mental *entities* open to empirical investigation. Both involve some revision of ordinary assumptions. The behaviourist denies that 'mental phenomena' ever cause or explain our behaviour. Similarly the causal analysts deny that causation *supports and explains* the regularities and counterfactuals that hold, since causation in fact *consists in* or *supervenes on* such regularities and counterfactuals.

In this light, the 'Process Theory' can be compared with a simple Type-Identity theory of mind, which suggests that (mental state M = physical state B). Sceptical of the possibility of repairing unsuccessful analyses, and aiming to respect the intuition that causation is something independent and explanatory of evidential relations, counterfactual relations and regularities, process theorists sketch a programmatic

model for an empirical identification of causation with some physically specifiable relation R. But there is an obvious problem here. For any proposed R, it seems clear that causation *might not have been* R; but as Kripke (1980) pointed out, all identities are necessary. Therefore, it cannot really be the case that (causation = R).

This is where Menzies (1996) approach comes in. Menzies proposes that we *combine* the two approaches to causation, as Lewis combined the two accounts of the mental, to solve the problems of both. Lewis's idea was that we could unite the insights of behaviourism -- that there are apriori truths about the mental to be has by conceptual analysis -- and the scientific respectability and ontological realism of the IDT.

First, he introduced a new method of analysis that could accommodate the holism of the mental with succumbing to circularity (see his `83a for details). Instead of attempting to reductively define the various types of mental state individually, we use Ryle-style conditionals as part of a long and detailed characterisation of a whole mental economy; the resulting theory could be presented as a complex sentence (S1) making free use of mental vocabulary ("is in pain,""believes that p" etc.) along with logical vocabulary and terms for causal relations. We then eliminate *all* the mental terms together in favour of existentially quantified variables (plus a uniqueness specification), giving **(S1′)** -- the Ramsey sentence of S1 -- saying that there are states $x1...xn$ which stand in this structure of relations (call it R) to each other, perceptual input, and behavioural output; and that there is no distinct set of states which stands in R. We can then specify what it is to be in any one of these mental states. Where "Mi" is the mental state term replaced by xi: to be in M1 is to be such that there are states satisfying **(S1′)** and one has x1.

However, we can now potentially say more about M1. For what we have in (S1) is a *descriptive specification* of the mental states M1...Mn; so if we discover that certain physical states uniquely satisfy that description, it seems that we can and must *identify* (Mi=Pi) (Compare: G.Orwell is whoever wrote *Animal Farm*, E.Blair wrote *Animal Farm*, therefore G.Orwell is E.Blair.) Concepts of mental properties have apriori analyses; mental properties have scientifically discoverable natures; the former determine the latter. Further, this newly grounded IDT can escape Kripke's

problem. Kripke's point is that all identities hold necessarily, and Lewis agrees. Nevertheless, Kripke allows contingently true identity-*statements* (where 'an identity statement' can be understood in terms of the surface structure of two singular terms flanking the "=" sign), and this is all that the IDT-ist requires. Just as we can consistently accept "Orwell is the F" ("Orwell uniquely instantiates F") together with "Orwell might not have been the F," so we can consistently say that pain is B, but pain might not have been B.

Now, Menzies recognises (a) that the Probabilistic theories of causation capture important, obviously correct generalisations about causation (that it in involves the raising of chances and counterfactual truths) but (b) that they nevertheless fail at crucial hurdles; and (c) that this is to be expected, given that our concept of causation is the concept of a relation that *underlies and explains* the more superficial connections appealed to in the analyses, which in turn *evidence* it.

The suggestion then is that we take seriously this aspect of the concept, by treating it as a *theoretical* concept, in the way that Lewis treats psychological concepts. Having given a descriptive characterisation of causation in this way, the stage would be set for the contingent identification proposed by the process theory.

Menzies suggests three "crucial platitudes" for the theoretical definition:

(P1) Causation is a relation which holds between *distinct* events.

The notion of 'distinctness' is no more sharply defined than we should expect, in the context; it reflects our recognition of various forms of dependence relations between events as *excluding* a causal connection. Examples of such relations are rife in Kim's work. There is part/whole dependence (e.g., of my writing "writing" on my writing "it"; the dependence of an extrinsically characterised event on an in intrinsically characterised event (e.g., of the widowing of Xanthippe on the death of Socrates); and relation of supervenience, realisation, and constitution.

(P2) Causation is an *intrinsic* relation between events.

This the principle which I have appealed to throughout as expressing the non-Humean aspect of our concept of causation: when two events are related as cause and

effect, there is some connection between them which is determined by the in intrinsic natures of. The exact content of this principle depends on how we give substance to the notion of *intrinsicness*, which is notoriously intractable. Menzies account appeals to Lewis's notion of 'natural' properties.[11]

First, we say that x and y are *duplicates iff* they share all their natural properties; and the pairs <x, y> and <x2,y2> are *duplicate pairs iff* x and x2 are duplicates, y and y2 are duplicates, and the natural relations holding between x and y are exactly those holding between x2 and y2. Then, a relation R is:

(I) intrinsic to its relata iff [(x)(x2)(y)(y2) x and x2 are duplicates & y and y2 are duplicates $\rightarrow$ (R(x,y)$\leftrightarrow$R(x2,y2))]

(ii) intrinsic to its pairs iff [for all duplicate pairs <x, y> and <x2,y2> (R(x,y) $\leftrightarrow$ R(x2,y2))]

(iii) extrinsic iff its not intrinsic according to the above.

(P3) Causation typically coincides with the raising of chances

In sum, then, we have the following model of theoretical definition of the causal relation: it is that intrinsic relation that typically holds between two events when one increases the chance of the other.

I will say more about Menzies' project after I have considered some proposed candidates for the realizer of the causation-role (the theories of Aronson and Fair, see below). First, I need to discuss Salmon's (`80,`84) process theory, which doe not quite fit into the scheme laid out so far. Although Salmon appeals to the notion of a process as the crucial core of causation his account involves the interaction and interdefinition of many different concepts and elements that it is easy to lose track of, despite his admirably informal exposition.

The two fundamental notions of the account are *production* and *propagation* (p.139).

(1) Causal processes propagate causal influence, in the form of propensities, throughout space and time. A '*process*' (after Russell 1948) is anything that displays consistency of *structure* over time. The key distinction is that between 'causal

---

[11] see Lewis 1983b. The details will not be important here.

processes' and 'pseudo-processes': a process is *causal* iff it 'transmits its own structure'. The empirical criterion of this feature is the ability (not necessarily exercised) to *transmit a mark* or local modification in structure. "Mark transmission" is defined by the 'At-At' principle, as follows. Consider a process P, a characteristic Q and distinct space-time points α and β such that, in the absence of interactions with other processes, P would remain uniform with respect to Q, manifesting it over an interval including both α and β. Let M be a mark consisting of a modification of Q into Q' introduced by a single local interaction at α. Then: M is transmitted from α to β iff P manifests Q' at β and at all stages of the process between α and β, without additional intervention. So, e.g., a ray of light travelling from a torch to a wall is a causal process; whereas if I swing the torch in a circle around me, the movement of the spot of light along the wall is a pseudo process.

There are two notions that Salmon associates with processes. The first has already been introduced. Following Russell's ('48) characterisation of a 'causal line', he claims that they each have a certain *structure*: the persistence or uniformity over time of certain qualities or characteristics (p.144). In multiplicity, they may exhibit *order* e.g., the kind of correlation that we might find in the paths of two shards of shrapnel from an exploded bomb.

(2) The structure of, and order amongst, processes is produced I situations characterised by three kinds of statistical 'fork', corresponding to the three kinds of common cause.

What's important is that the existence of such forks does not guarantee a causal connection: there must be appropriate *processes* connecting the events.

These 'events' - the relata of causation - are intersections of such processes. Salmon characterises the case in which such an intersection constitutes a *causal interaction* as follows. suppose P1 and P2 are two processes that intersect at space time point s. Q is a characteristic that P1 would exhibit throughout an interval around s, if the intersection with P2 did not occur; R is a characteristic that P2 would exhibit throughout an interval around s, if the intersection with P1 did not occur. Then, the intersection constitutes a causal interaction if:

(a) P1 exhibits Q before s, But exhibits a modified characteristic Q' throughout an interval immediately following s

(b) P2 exhibits R before s, but exhibits a modified characteristic R' throughout an interval immediately following s (p.171).

So for example, my hitting a ball with a racquet is a causal interaction: two 'lines' (the swinging racquet and the stationary ball) intersect, with a modification in the characteristics of each. Hitting (C) forms a fork with striking the opposite end of the court (A) and subsequently striking the perimeter fence (B). But the fork is a causal one only in virtue of there being processes running from this instance of C to this instance of B, and similarly for C and A. No pre-empted cause will count as a real cause, because there is no causal process connecting it to the effect.

There are two questions about Salmon's theory that I want to ask. The first is whether it avoids circularity; the second, concerns its use of the notion of 'process'.

The suggestive and interdefined nature of the terminology brings an air of circularity to the theory, but we must be careful here: the question is whether he successfully eliminates causal presuppositions in his explication of that terminology. For instance, Mellor (ˆ91) charges Salmon with circularity on the grounds that his theory uses the notion of 'production', which is just a synonym of 'causation'. But it would be uncharitable to suppose that salmon would fall victim to such flagrant triviality (his ˆ80 is *titled* "Causality: Production and Propagation"); and on closer inspection, we find that the notion of production is actually cashed out in the theory in terms of the marking of processes in interactions. We have production where we have **(1)** a causal interaction (i.e., an intersection of two processes - see above) in which**(2)** both processes are 'marked', a mark being **(3)** a modification or alteration to a characteristic introduced by a single local interaction. Any circularity must be found amongst these notions. So, as it stands, Mellor's accusation seems to be no more just than one directed at his own account's appeal to causes "raising the chances of" or "giving a certain chance to" their effects. These locutions simply have a causal import in ordinary usage that is not presupposed in the philosophical relations (i.e.,probabilistic counterfactual dependence) which they serve to introduce.

However, there are deeper circles to be found. One is that (3), above, used in characterising 'production', involves the idea of *introducing* a change in a characteristic. For this to be acceptable, we must suppose that Salmon intends no more than Humean sequence: the intersection is followed by a change in characteristic. More worrying is the use of "interaction" in (3). Suppose that such a Mark introducing interaction could be constituted by the intersection of two 'pseudo-processes'(e.g., moving shadows or spots of projected light.) Then a pseudo-interaction could be a cause. But the whole point of Salmon's distinction between causal and pseudo processes is that the latter cannot transmit causal influence. So, the only interactions that introduce marks are causal interactions. This means that the notion of 'marking' presupposes that of 'a causal interaction' and so cannot be used to explain it (that is: our understanding of (3) presupposes our understanding of (1) and cannot be used to explain it .) This is clearly unsatisfactory.

What about the notion of a process? Clearly, this is the fulcrum of the theory (he explicitly claims that it provide the missing connection sought by Hume); and this only emphasises the fact that it is a technical philosophical notion that cannot be taken for granted. After all, it is not very illuminating to tell us that pre-empting causes are distinguished from pre-empted cause by a causal process unless we have *some* kind of characterisation of a process. Lewis says as much; he just doesn't think that it explains anything.

How do we get a hold on processes ?

As with facts and events, we can't get very far on the basis of how the word "process" explicitly occurs in ordinary usage. Admittedly, there are plenty of things which we would *not* call processes, for example clicks, knocks, storms, physical objects, societies; but it isn't obvious what they all lack which excludes them from process-status (if we can't find a unifying feature of all *games*, we're not likely to find any interesting account of *things that aren' t games*). It can't simply be the absence of (sufficient) temporal duration; human beings persist, and football matches have substantial temporal duration, but we don't ordinarily call either kind of thing a process. Perhaps processes occupy time in a different way to such things: processes are more naturally said to go on for a period. But, turning to our positive application

of "process", this doesn't look very promising.   Similarly we might say "filing one's invoices is a tedious process", "persuading him to come to the conference was a laborious process." It is transparently obvious that there is nothing ontologically significant about these 'processes'. The common element seems to be some *goal* or *purpose*, in each case, that gives teleological structure to a collection of events. for example, in baking a cake, one does various things, with some degree of systematicity and order, that are required if one is to achieve the end product, viz. a cake. That makes the development of the events that make up the whole event a *process*. Clearly, there is no further thing *connecting* the sub-events that we might pick out with "process"(although there are certainly intentions and desires that might do the job - see below)

Nevertheless, A.P.D Mourelatos (´78) has tried to draw an ontological distinction on the basis of some subtler linguistic considerations. His line of thought begins in response to Vendler's (´57) and Kenny's (´63) closely related attempts to categorise verb-types on the basis of data concerning how tense affects implication relations. We can get the flavour of the project by considering some examples (I follow Vendler's typology).

(a) "John's pushing the cart uphill"

(b) "John's climbing Ben Nevis"

(c) "John's reaching the summit"

(d) "John's knowing the safest way down"

(a) is an *activity*: it is homogeneous - its parts satisfy the same description as the activity itself - and has no natural terminus or culmination. **(b)** is an *accomplishment*: it may have a culmination, but it cannot be homogeneous. **(c)** is an *achievement:* it is a culmination, and lacks temporal duration. Finally, **(d)** is a *state:* it is homogeneous like (a), it persists or endures, but we cannot sensibly ask how long it took, or whether it culminated. As Mourelatos shows, such classifications don't quite come off, because many of these verbs are 'multivalent', satisfying the criteria for different categories depending on the context of occurrence. For instance, "know" can sometimes occur as an achievement word ("and suddenly, I knew what I had to do"); "understand" can occur as an activity-word ("I'm understanding more

about quantum theory every day".) These and other considerations indicate that such distinctions need to be drawn at the level of whole predications, and are importantly governed by the phenomenon of *aspect:* 'perfective' or 'imperfective'(for more on aspect, see Mourelatos; Smith `91)

Mourelatos notes a certain parallel between the familiar count/mass distinction in talk about objects, and differences of character amongst (what he calls) the 'nominalization transcriptions' of verb-al predications in talk about occurrences (events, situations etc.) On the basis of this, he draws an analogy between the makeshift genera of "things" (e.g., particulars and stuffs) and "eventualities"(e.g., events), with respect ot which he apples two contrasts strongly associated with aspect, to draw an ontological distinction between processes and events: (i) Homogeneity vs. Heterogeneity and (ii) Direct countability vs. Indirect countability.

The count/mass distinction concerns (certain uses of) nouns. 'Count nouns'(e.g., 'dog', 'boat', 'tree') take the indefinite article and pluralisation; they accept numerals and count-modifiers; and they can be governed by an informal existential quantifier "there is at least one..." Mass terms (e.g., "wine", "snow", "gold") accept these transformations, if at all, only with a change in their grammatical role; and they have their own characteristic quantificational expressions. So, for example, "We stock over 100 wines" means that they stock over 100 *kinds of* wine; and we can say "do you think we have enough/too much wine?"

Nominalisation-transcription is a generalisation of the method (see (II)) of transforming a sentence describing (in Kim's sense) an occurrence in to a sentence quantifying over occurrences. We have already seen the nominalisation transcription of sentences about events, such as the move from **EP** to **ET** :

[EP] Jones capsized the canoe

[ET] There was a capsizing of the canoe by Jones

Mourelatos suggests that the same features that characterise the count/mass distinction mark a distinction between the transcriptions of event and process predications. So, consider the following examples:

[PP1] "Jones pushed the cart for hours"

[PT1] "For hours there was pushing of the cart by Jones"

[PP2] Jones was painting the nativity

[PT2] There was painting of the nativity by Jones

These transcriptions are *mass-quantified.* There is no terminus or closure of the activities specified in **PP1** and **PP2** that would allow for an indefinite article in **PT1** and **PT2**. Numerals and (cardinal) count adverbials are out of place in such predications or their transcriptions. The conclusion is a criterion for event predications: they are all and only those that have count-quantified nominalisation transcriptions.

Mourelatos then characterises the ontological categories of 'process' and 'event' in terms of two distinctions. **(i)** Events, like objects, are heterogeneous, whereas processes are homogeneous or 'homoeomerous'. So, just as houses don't standardly have houses as parts, "I built a house" is not true of any subintervals of the time of the occurrence which it applies to; whereas just as "snow" applies to every subsection of a quantity of snow, so "John is running" is true of every subinterval of a period in which John is running. **(ii)** Events and objects are directly countable whereas stuffs and processes are inherently indefinite. So, just as we can ask "how many dogs?" we can ask "how many storms ?", since events have terminal points that mark their completion, dividing them into distinct units. By contrast, there is no culmination or anticipated result of processes, in the same way that stuffs fail to divide into natural units; we can't ask "how many snowings ?"any more than we can ask "how many snows?" They can be counted only by reference to possible spatial or temporal containers: "he ran three laps", "three pounds of gold".

Now the linguistic considerations noted by Mourelatos have been recognised by others (Taylor `85 for a very similar account with some formal additions; Link `87 for sophisticated developments) who have not followed him in drawing ontological distinctions. And there is good reason to be cautious about such a step.

First, consider 'homogeneity'. As Taylor notes, the putative homogeneity of certain occurrences breaks down on closer inspection. For example, chuckling or rumbling is only relatively homogeneous: no 1/10 second of sound counts as chuckling. Rather, there are minimal periods of chuckling which we can identify as

'those which everyday empirical criteria can identify as such'(similarly stuffs: it is not true that every subregion of the location of a quantity of gold (let alone fruitcake)contain the same stuff.) Hence, linguistic homogeneity of process-predication, basically characterised by the entailment from "X is F-ing" to "X has F'd" must also be a matter of degree, allowing for a first minimum period within which the former will hold but the latter will not (Taylor gives a series of postulates explicating the logical properties of the various sorts of predicate)

A sceptical line of thought here is that the linguistic homogeneity that is a matter of how predicates apply to temporal intervals and subintervals isn't' a reflection of a metaphysical homogeneity that consists in empirically observable structural similarity (cf. Salmon's notion of 'structure' above). For example, "the regeneration of the British film industry" would seem to count, semantically, as picking out a process: any subinterval of a period in which the British film industry regenerates is a period of the British film industry's regenerating; but there is no appearance of metaphysical homogeneity here (similarly "the depopulation of the rainforests", "the evolution of Homo Sapiens".) Rather it has to do with our seeing certain series of events as developing towards a conclusion and/or developing from one to the next in a causal way. Hence we so often regard purposive sequences of actions as processes: making a cake, or building a model ship. Mental processes and the 'creative process' are paradigms, and in the non-mental world, growth or healing; these are tied together by the idea of an ordered development from one stage to the next. But the reality of such order amongst events, while undoubtedly important, does not suggest a new kind of entity. No-one should be considered a sceptic about the reality of 'cognitive dynamics' or the structure of intention on the grounds that she posits only events, states, and dispositions.

What about the contrast of 'direct' and 'indirect' countability? I think that a similar problem arises here. Suppose that Jones pushes the cart to the top of the hill, taking two hours about it (and he doesn't push it anywhere else.) Then we can truly say:

(1) Jones pushed the cart for two hours   and

(2) Jones pushed the cart to the top of the hill.

By Mourelatos criteria, we can discern a process in (1) ("there was pushing...") and an event in (2) ("there was a pushing..."). So we seem to have here two coinciding entices of different kinds. Mourelatos would, I think, in line with his analogy with objects and stuffs, treat this like the case of a ring made up of a quantity of gold: a pushing-event made up of a certain amount of pushing-stuff. The event is constituted out of process. Now, we should already be sceptical about drawing a serious ontological distinction on this basis. For when we look at what actually happens in the world, we may find that Jones progress to the top of the hill was a staggered, disjointed one; maybe he stopped and started , so there is no appearance of a quantity of homogeneous pushing stuff. Further, suppose that Jones read aloud the letter "A" this morning (and nothing else). Parallel with the above, we have:

(3) Jones read

(4) Jones read the letter "A"

again, (3) counts as picking out a process, (4) an event. But the observable facts completely undermine the idea that her reading of the letter was made up of some ontologically distinct entity, a quantity of reading. More generally, consider cases like these:

(5) Jones irritated me yesterday

(5') Jones was irritating me yesterday

(6) Jones looks good

(6') Jones is looking good

In both 5' and 6', we achieve a certain effect via imperfective aspect marked here by the progressive form (these points are made by e.g. Smith `91.) The happening described by 5 and the state described by 6 are presented from an 'internal' point of view (in 5') or as 'dynamic' (in 6'); it seems clear that no ontological distinction is warranted.

In conclusion of this discussion then, there is no sound basis for a notion of process here. One potentially important candidate has been suggested: a structural aspect of reality, concerning how events are ordered or governed by purpose or

teleology. But obviously, this is of no use for the purposes of explaining causation, not least because this kind of process plausibly *presupposes* causal connections.

Of course, Salmon would not appeal to such considerations to characterise the notion of a process; but he says very little, and I have pressed the point to establish that such a characterisation needs to be given. Anyway, I suggest that we give up on any general or abstract account of processes, in favour of some concrete proposals.

This is what Salmon tells us about processes:- **(i)** Stationary or moving particular material objects are processes, as are rays of light, waves of sound, and radio signals **(ii)** Processes have a 'structure' or 'uniformity of characteristics' that is closely related to the transmission of energy (pp.154-5) or information (p.261) **(iii)** All and only causal processes transmit energy (p.146). I think that (ii) and (iii) give us the more general notion that Salmon has in mind; he takes physical objects to be a particular form of process that affords us some intuitive concrete examples of observable causal links. So I now turn to some more specific process theories.

These theories identify the process as the *transference of energy or momentum*: something that persists the same across space and time (just as a middle sized physical object does); this is the idea suggested by Salmon's comments above. Aronson's (´71) analysis is as follows:

[C1] In "a causes b", "b" designates a *change* in an object that is an *unnatural* change (see below)

[C2] In "a causes b", at the time b occurs, the object that causes b is in contact with the object that undergoes the change

[C3] Prior to the time of occurrence of b, the body that makes contact with the effect object *possesses a quantity* (e.g., heat, kinetic energy, momentum, velocity), which is *transferred to* the effect object (when contact is made) and *manifested as b*.

Aronson's piece is noteworthy for recognising the significance of transitive verbs for the theory of causation (as does Anscombe ´71). Aronson, in the spirit of Austin, regards "cause" as 'dimension-word', serving as " the most comprehensive term in a whole group of the same kind that fulfil the same function"; and he explicitly says that "a full understanding of causation would require an understanding of the

meaning of those [i.e.,transitive] verbs". Forgoing this task, he claims that his analysis offers transference as the underpinning for "mechanical transitive verbs."

It is pretty obvious that Aronson's account puts serious restrictions on bona fide cases of causation. But we can find difficulties for his key condition [C3] (Aronson finds two himself, (iv) and (v) below) without departing from the physical/mechanical context that he has in mind.

(i) We blow carbon dioxide onto a candle (or just put a glass over it), extinguishing the flame - causing it to go out. Neither the gas nor the glass would transfer anything to the flame in causing it to go out.

(ii) I pull the plug out of the bath at t1, causing it to be empty at t2. There's no apparent transference here.

(iii) I pull the plug out of the socket, causing the light go off. The removal of the plug doesn't transfer any energy; it prevents or curtails such a transfer.

(iv) We chill water (i.e.) cause it to become cooler by putting ice in it, or putting it in the refrigerator. But the transference runs the wrong way here: the ice absorbs energy from the water.

(v) A weight is attached to a spring and the spring stretched; the weight is then held in place with a catch. Releasing the spring will cause the weight to accelerate, but it will not transfer anything to the weight.

Clearly, none of these are peculiar cases; but they are problematic for Aronson's theory.

Aronson takes a revisionary line with (iv) and v) which he would presumably extend to the other counter examples. His response to (iv) seems to be that it merely reflects our mistaken tendency to think of objects that absorb heat as having something, 'coldness', which they emit or transfer. what really happens, as we know, is that heat is transferred to the ice, from the water, and is converted into kinetic energy which overcomes the forces bonding its molecules; "as a result, but not an effect, of this process, the water loses its heat and thus becomes cooler." The real description of the case is that in causing the ice to melt, the water gives up some of its heat i.e., becomes cooler.

This is blatantly stipulative and unconvincing. We can say, with perfect coherence, "Of course, the ice doesn't transfer its 'coldness' to the water; it causes the water to become cooler by absorbing heat energy from it." The apparently reasonable claim that the ice causes the water to become cooler thus does not rest on any confused thinking about transference. It's also true that the water causes the ice to melt; but that doesn't tell against (iv) as a distinct true claim. There's nothing odd about two causal connections here, unless we assume that causation is transference.

Aronson treats (v) as an instance of the general phenomenon of 'triggering.' What happens is that when we release the weight, the spring transfers its potential energy to the weight, which manifests this quantity as kinetic energy in the acceleration that is the effect. So, we make a general distinction between "causes" and "occasions", the former being things that bring about the changes, the latter conditions for making V-ing possible - "that which enables a cause to act." Releasing the weight is not a cause, but an occasion.

However, this is an implausible distinction that has never been substantiated. Admittedly, there is a nice question (see e.g. Mackie's `65) about the principles underlying the distinctions we undoubtedly make between causes and 'standing conditions', but it is a question about the pragmatics of causal discourse, and does not concern what can and cannot be a genuine cause. The fact is: in the majority of causal situations, we have a range of contributing factors, each depending on the others for its efficacy. So each makes it possible for the others to act, but this does not of course mean that we are left with no causes. A familiar example: a spark, a quantity of inflammable liquid and the presence of oxygen will combine to yield an explosion. When we start a car, we properly regard the spark as the cause; if there is an accidental explosion in a factory, we will with equal propriety regard the the carelessly stowed petrol as the cause. Nevertheless, it is the same situation in both cases.

We might also question the necessity of condition **C1**, on which **C3** depends. The distinction between 'natural' and 'unnatural' changes is that between changes that take place independently of other objects - for example, a body's continuing to move in a straight line with a constant velocity - and changes that result from

interactions with other objects: all and only the latter, according to Aronson, are caused. But this doesn't seem obvious. For instance, the decay of a quantity of radioactive matter appealed to in various places by Salmon, could be a causal sequence. More familiarly, the process by which sugar turns to alcohol in the production of wine is a causal one, which also does not seem to involve the transference of energy between two distinct objects (although energy is generated in the process). It would be stipulative to insist on two objects in every case of causation.

Aronson's analysis, then, is inadequate, in fundamental ways (The same goes for Castaneda's `84 proposal, which posits the transfer of some unknown quantity 'causity' as the basis of causation.) There are several other problems, but these can be considered in the context of Fair's (`79) refined analysis, to which I now turn.

The core is effectively as follows. **a** is a cause of **b** iff there are physical redescriptions of a and b as manifestations of energy or momentum, such that one of **(F1-F4)** holds:

[**F1**] this energy or momentum is, at least in part, transferred from a-objects to b-objects (p.243)

[**F2**] this energy or momentum is transferred between a-objects and b-objects, and the redescription of a is a description of an energy gain, the redescription of b the description of an energy loss (p.243)

[**F3**] the redescription of a is of a lowering of a barrier to the release of potential energy, an d the redescription of b is of a manifestation of that potential energy

[**F4**] the redescription of a is of the raising of a barrier to the flow of energy, and b has a redescription as the interruption of the flow of energy. (pp.244-5)

On the face of it, this is inappropriately ad-hoc as a theory of causation. For it avoids the problems faced by Aronson simply by adding escape-clauses ((F2) handles (iv); (F3) accommodates iii and v; (F4) deals with i and ii; and Fair goes on to add three further condition s to cope with three ways in which omissions can be involved in causation.) The danger is that we are just locating a variety of phenomena that account for a variety of cases of causation; and that isn't really illuminating. After all, a lot, perhaps the majority, of macroscopic causal interactions in our world

involve inanimate physical entities bumping and deflecting each other; and as things are, every such transaction can be explained in terms of some form of energy transfer. Fair wants to say more than that:

"the hypothesised relationship between causation and energy flows is expected to have the logical status of an empirically discoverable identity, namely that causation is identical with a certain physically specifiable relation" - the proper specification waiting on the development of a unified science.

From this point of view, Fair's account doesn't look promising. Why think that there is *any* interesting physically specifiable correlate of causation ?

Here's a deeper problem for both Fair and Aronson. The process theory is promising, from the point of view of problems like pre-emption and determining causal direction, because it posits some thing that can be identified at different times at different locations. We can tell, say, which of two simultaneous blows caused the death by tracing back the momentum that jarred the brain to one rather than the other. But are quantities of energy things that can be numerically identical over time? What is the difference, for instance, between a number of particles sequentially exhibiting the dispositions definitive of a certain charge-property, and there being a genuine 'transference' from one to the other? More generally, how do we identify quantities of energy?

Fair claims that in a closed system, conservation principles allow us to reduce the problem of identifying quantities of energy or momentum to that of identifying objects. Fair uses the model of a sequence of collisions between billiard balls: **a** hits **b** causing it to move and hit **c**, which moves itself; **a** thus causes **c** to move. Prior to it's contact with **b**, **a** is a closed system, in which energy and momentum are conserved. We can identify the energy lost by **a** in the collision with **b** (minus losses to friction) with the energy and momentum gained by **b**; and similarly for **b** and **c**. " The theoretical apparatus is sufficient to determine that **b** transfers the same energy, or part thereof, to **c** as it acquired in the collision with **a**." But as Ehring (1985) points out, this isn't generally so. For suppose that **a** gives up part of it energy to **b** retaining some itself; similarly, **b** gives up part of its energy to **c** retaining some itself. How do

we tell whether that part of **b**'s energy that is passed on to **c** is the very same energy that **b** received from **a** ?

A second deep problem for these process theories arises with respect to mental causation. On one hand, there is a specific problem about mental causes and the flow of energy. On the other, there is a more general problem of heterogeneity, which the mental merely provides obvious instances of: causes and effects *aren't* always manifestations of energy, and when they are not, we tend to find that there is no appropriate transfer of energy to identify as the causal connection.

So, first, given that causation just *is* a certain physical relation - say, the transference of energy, ignoring the various qualifications - the ascription of a causal role to mental features of the world could be literally correct only if there is a physical flow from those features to other aspects (mental or physical.) This, in turn, would seem to present us with just two options for saving mental causation:
(1) Mental features are not identical with physical features, yet they have causal efficacy by contributing psychologically generated energy, distinct from any physically sourced energy - to the world[12].

This is not, I suppose, an hypothesis too absurd to countenance (it might be true for all I know). But I don't think it's the right story for the realist about mental causation to stake his claim on. I think it concedes too much to the physicalist conception from the start.
(2) An Identity theory: the causal flow form mental features is just another physical flow from physical features, because the mental features *are* (identical with) physical features.

I can't evaluate identity theories here. All I need is the following point: even if such a theory were true, it would not be enough to secure a causal role for the mental. Thus far, I have been talking loosely of 'mental features'; but this disguises significant distinctions amongst the aspects of the psychological that we are interested in. The paradigm case for the process theory is the sudden pain, when, say, one touches a hot surface. The idea would be that the instance of pain be identified with

---

[12] see W.D.Hart 1988 for something along these lines.

the instantiation of a certain property in the brain, the causal efficacy of the pain being located in the electrical impulses travelling from the central nervous system to receptors in the muscles, and so on. Unfortunately, even if one finds that plausible, it isn't a very good paradigm for thinking about mental causation; such situations are really relatively unusual. For a start, nearly everything we do is governed by what beliefs and 'desires'('pro-attitudes') we have: whenever I take the bus, or pick something off a supermarket shelf, or whatever, I do so on the basis of what I want, and what I believe about how to achieve it (this is oversimplified, but there's enough truth in it to make the point I'm after). Such attitudes are not events -- punctual occurrences -- but states or dispositions. Admittedly, philosophers sometimes talk of "occurrent beliefs", which we might best construe in terms of thoughts which manifest beliefs; and there is often such a conscious thought prior to and during intentional action. Nevertheless, it is plain wrong to suppose that it is *only* such conscious happenings that have a causal role. Consider an uncontroversial case of a state: knowing the way to the supermarket. We can cite my knowledge of the route in explanation of my going that way to the supermarket, even though I may do so quite 'unthinkingly'. This is a genuine - causal - explanation in virtue of the fact that I move as I do because of my cognitive state; it is genuinely rational (rationalising) explanation because that knowledge that guides my behaviour is accessible to rational reflection on my part, and manifestible in other ways (for instance, I can, other things being equal, tell *you* how to get to the same place.) I take it that to be obvious that there is no hope of construing the causal role of my knowledge in terms of energy transfer, even if it is assumed to be identifiable in terms of syntactic form and functional role. The problematic examples are innumerable: she stays with him because she loves him; she has been angry for the last fifteen years because of the way the miners were treated under the Tories; he apologised because he was nervous.

This only scratches the surface of the issue; mental causation deserves a fuller treatment than I can give it here. However, the two problems should be sufficiently clear. First, there is a serious problem allowing a distinct role to non-physical, mental features if causation is identified with energy transfer. This will not carry as much suasive force as I think it should, since the idea that there is a distinct causal

role for anything non-physical is unpopular anyway. But regardless of this, mental causation shows that process theories of causation are generally just too restrictive about what can count as genuine causal relata, and correspondingly about what can count as causation. There are plenty of non-mental examples: We say that the slump caused the fall in the price of gold, that the snow on the line caused the train to arrive two hours late, that my nose is the shape it is because of my genetic inheritance. The causes and effects here are *not* manifestations of energy, and the causal connections are not transfers of energy.

Menzies, since he is not advancing any particular candidate to be identified as a causal relation is not directly undermined by the counterexamples that derail Fair and Aronson's theories; but their failure should make us wonder what else *could* possibly play the role.

CONCLUSION

I have reviewed a range of theories that attempt to explain or reduce causation and found them doubly inadequate:

(i) they all artificially restrict the subject-matter

(ii) none achieves its explanatory ambitions, even within its chosen limits

I suggest that if we attend to the range and variety that we actually find under the head "causation" - rather than starting from paradigms which are really not representative - this project would appear at best unlikely to succeed, at worst misguided.

I think that Menzies' assumption that we have (or can distil) a single, general concept of causation, which is effectively a (non-rigid) descriptive charaterization of an *entity* - a relation - is a misrepresentation. Rather, we have a range of causal concepts, which express different kinds of production and dependence (we may call these 'relations'if we wish; but positing entities does not here explain anything). We can group these together under the philosophical term 'causation' - but this does not show that there is a common concept or a common relation that is open to a general explanation.

Further, some of these 'causal relations' are bound to remain unexplained; in particular, those expressed by transitive verb constructions of the form [NP VP N'] or [NP VP [N PP N']] (e.g., "John knocked the vase onto the floor"). How might Menzies' incorporate such cases into a theory of a relation between distinct events ? What, I suspect, Menzies intends is that such constructions be analysed into the form "A caused (it to be the case that) B" -- where "A" is an expression for the causing state of affairs, "B" an expression for some end-state that is brought about -- and these expansions reflected in the theoretical definition.

Now there is no doubt that the verb constructions entail such expansions. For instance, if John closed the door, then the door is closed and John did something that brought that state of affairs about. But that something might be nothing less than *closing* it. The expansion, notice, does not ensure the truth of the original claim, because of the familiar possibility of 'deviant causal chains': there are various ways in

which something John does results in the door's being closed that don't count as cases of him closing the door. This is a quite general problem, and there is no general solution. I move the knife, the butter ends up in two, and my moving the knife caused the butter to end up in two. It doesn't follow that the knife *cut* the butter; maybe my movements generated a freak burst of heat that dissected the butter, followed at every instant by the blade. The expansions aren't equivalent to the originals.

This seems to me to be an instance of the general failure of 'componential' analyses of basic notions (compare e.g. Davidson's attempts with 'acting for a reason'(`80e see esp. essays 1-5); Snowdon (`81, 90) on "x sees y"; Williamson `95 on knowledge). In each case, we have a 'starting condition', a 'terminal condition', a causal condition, and an intractable fourth requirement that is required to ensure that the causal connection is of the 'appropriate' kind. The conclusion is that the analysandum is at least as basic as the notions appealed to in the analysans; and this I suspect is how the various concepts expressed by transitive verbs stand with respect to the concept 'causes.'

BIBLIOGRAPHY

Anscombe, G.E.M. (1971) *Causality and Determination* reprinted in Sosa and Tooley (eds.)

Aronson, J.L. (1971) 'On the grammar of "cause"' in *Synthese* 22: 414-430

Barwise, J. and Perry, J. (1981) 'Semantic Innocence and Uncompromising Situations'in French, Uehling, and Wettstein (eds.) *Midwest Studies in Philosophy VI.* Minneapolis: University of Minnesota Press

Barwise,J. and Perry,J. (1983) *Situations and Attitudes.* Cambridge MA: MIT Press

Beebee, H.(1998) 'Do causes raise the chances of effects ?' in Analysis 58.3, pp.182-90

Bennett, J. (1989) *Events and their names.* Cambridge: Cambridge University Press

Cartwright, N. (1983) *How the laws of physics lie.* Oxford: Clarendon Press

Castaneda, H.N.(1984) 'Causes, Causity, and Energy' in French, Uehling, and Wettstein (eds.) *Midwest Studies in Philosophy IX.* Minneapolis: University of Minnesota Press

Chalmers, D.J. (1996) *The Conscious Mind: In Search Of A Fundamental Theory.* Oxford: Oxford University Press.

Chomsky, N. (1975) 'Some Remarks on Nominalisation' in Davidson and Harman (eds.) *The Logic of Grammar.* Dickenson: Encino California.

Davidson, D. (1980a) 'Causal relations' repr. in his 1980e

Davidson, D. (1980b) 'The individuation of events' repr. in his 1980e

Davidson, D. (1980c) 'Events as particulars' repr. in his 1980e

Davidson, D. (1980d) 'Mental Events' repr. in his 1980e

Davidson, D. (1980e) *Essays on actions and events.* Oxford: Clarendon Press

Donnellan, K.(1966) 'Reference and Definite Descriptions' in *The Philosophical Review* 77: 281-304

Ducasse, C.J. (1926) 'On the nature and observability of the causal relation' repr. in Sosa and Tooley (eds.) 1993

Eells, E (1991) *Probabilistic Causality.* Cambridge: Cambridge Universiry Press

Eells,E. and Sober, E. (1983) 'Probabilistic Causality and the Problem of Transitivity' in *Philosophy of Science* 50: 35-57

Ehring, D (1985) 'The transference theory of causation' in *Synthese* 67: 249-258

Fair, D. (1979) 'Causation and the flow of energy' in *Erkenntnis* 14: 219-250

Fales, E (1990) *Causation and Universals*. London: Routledge

Hart, W.D.(1988) *The engines of the soul*. Cambridge: Cambridge University Press

Higginbotham, J. (1983) 'The logic of perceptual reports' in *The Journal of Philosophy* 80: 100-127

Jackson, F.(1994) 'Armchair Metaphysics' in J.O'Leary-Hawthorne and M.Michael (eds.) *Philosophy in Mind*. Dordrecht: Kluwer

Kenny, A.(1963) *Action, Emotion and Will*. London: Routledge and Kegan-Paul

Kim, J.(1966) 'On the psycho-physical identity theory' in *American Philosophical Quarterly 3*: 227-35

Kim, J. (1969) 'Events and their descriptions' in N.Rescher (ed.) *Essays in honor of Carl G. Hempel*. Dordrecht :Reidel pp. 198-215

Kim, J. (1976) 'Events as property-exemplifications' repr. in his (1993)

Kim, J. (1993) *Supervenience and Mind*. Cambridge: Cambridge University Press

Kripke, S. (1980) *Naming and Necessity*. Oxford: Blackwell

Lewis, D. (1979) 'Counterfactual Dependence and Time's Arrow' in *Nous* 13: 455-476. repr in his 1986d

Lewis, D. (1983a) 'New work for a theory of universals' in *The Australasian Journal of Philosophy* 61: 343-77

Lewis, D.(1983b) 'How to define theoretical terms' in his 1983c

Lewis, D.(1983c) *Philosophical Papers* Vol.1. New York: Oxford University Pres

Lewis, D. (1986a) 'Causal Explanation' repr in his 1986d

Lewis, D. (1986b) 'Causation' and 'Postscripts' repr. in his 1986d

Lewis, D. (1986c) 'Events' repr. in his 1986d

Lewis, D.(1986d) *Philosophical Papers* Vol.2. New York: Oxford University Press

Link, E (1987) 'Algebraic semantics of event structures' in J.Groenendijk, M.Stokhof, and F.Veltman (eds.) Proceeding of the Amsterdam Colloquium. University of Amsterdam: pp.243 - 61

Mackie, J.L. (1965) 'Causes and conditions'in *American Philosophical Quarterly* 2: 245-264

Mellor, D.H. (1991) 'The singularly affecting facts of causation' repr. in his 1991b

Mellor, D.H. (1991b) *Matters of metaphysics*. Cambridge: Cambridge University Press

Mellor, D.H. (1995) *The Facts of causation*. London: Routledge

Menzies, P. (1989) ' A unified theory of causal relata' in *The Australasian Journal of Philosophy* (67) pp. 59-82

Menzies, P. (1996) 'Probabilistic causation and the pre-emption problem" in *Mind* 105: 85-117

Mourelatos, A.P.D. (1978) 'Events, Processes and States' in *Linguistics and Philosophy* 2: 415-434

Rosen, D.A (1978) 'In Defense Of A Probabilistic Theory Of Causality' in *Philosophy of Science* 45: 368-386

Russell, B. (1948) *Human Knowledge: Its Scope and Limits*. New York: Simon and Schuster

Salmon, W.C. (1980) 'Probabilistic Causality' repr. in Sosa and Tooley (eds.)1993

Salmon. W.C. (1984) *Scientific Explanation and the Causal Structure of the World*. Princeton NJ: Princeton University Press

Smith, C, (1991) *The Parameter of Aspect*. Boston, MA: Kluwer Academic Publishers

Snowdon, P. (1981) 'Perception, vision and causation' in *Proceedings of th Aristotelian Society* 81: 175-192

Snowdon, P. (1990) 'The Objects of Perception' in *Proceedings of the Aristotelian Society* Supplementary vol. 64: 121-150

Snowdon, P (1992) 'How To Interpret "Direct Perception" in Crane(ed.) *The Contents Of Experience*. Cambridge: Cambridge University Press

Sosa and Tooley (eds.) (1993) *Causation*. Oxford University Press

Suppes, P. (1984) *Probabilistic Metaphysics*. Oxford: Basil Blackwell

Suppes, P.(1970) *A Probabilistic Theory of Causality*. Amsterdam: North Holland

Taylor, B. (1985) *Modes of Occurrence: Verbs, Adverbs, and Events*. Basil Blackwell: Oxford.

Tooley, M. (1990a) 'The Nature of Causation: A Singularist Account' in D.Copp (ed.) *Canadian Philosophers. Canadian Journal of Philosophy* suppl. 16: 271- 322

Tooley, M (1990b) 'Causation: Reductionism versus Realism' in *Philosophy and Phenomenological Research* 50, Supplement: 215-36); also repr. in Sosa and Tooley `93

Vendler, Z (1957a) 'Facts and Events'in his `57b: pp. 122-146

Vendler, Z. (1957b) *Linguistics in Philosophy*. Cornell University Press: Ithaca

Williamson, T (1995) 'Is Knowing A State Of Mind ?' in *Mind* 104: 533-565

Yablo, S. (1992) 'Cause and Essence' in *Synthese* 93: 403-449

Yablo, S. (1992b) 'Mental Causation' in *The Philosophical Review* 101: 245-280