I. KVART

THE COUNTERFACTUAL ANALYSIS OF CAUSE

In this paper I first go through an argument against cause transitivity, then sketch the theories of probabilistic cause and counterfactuals that I advanced.[1] I then proceed to the main goal of this paper: to present the counterfactual analysis of cause that follows from them.

1. THE NON-TRANSITIVITY OF CAUSE

David Lewis's counterfactual analysis of cause consisted of the counterfactual conditional closed under transitivity.[2] Namely, a sufficient condition for A's being a cause of C is that $\sim A > \sim C$ be true; and a necessary as well as sufficient condition is that there be a series of true counterfactuals $\sim A > \sim E_1, \sim E_1 > \sim E_2, \ldots, \sim E_n > \sim C$ $(n > 0)$.

At the core of this analysis, cause is taken to be transitive; namely, if A is a cause of B and B of C, then A is a cause of C. I have argued, however, that cause is not transitive,[3] which, if correct, undermines Lewis's analysis.[4] Specifically, consider the following counter-example to the transitivity of cause:

EXAMPLE 1. A worker x was injured in a work-related injury. Let A be:

x's finger was cut off at t_1 by a machine.

His co-workers rushed him and the detached digit to a hospital, where an expert surgeon was on alert; let B be:

The surgeon connected the finger to x's hand at t_2.

The surgery was a great success, and indeed, C:

x's finger was functional at t_3.

Synthese **127:** 389–427, 2001.

A, B and C occurred at t_1, t_2 and t_3 respectively. Intuitively, A was *not* a cause of C – the fact that x's finger was cut off at t_1 was not a cause of the fact that x's finger was functional at t_3.[5] But, again intuitively, A *was* a cause of B – the fact that x's finger was cut off at t_1 was a cause of the fact that the surgeon connected the finger at t_2; and, again, intuitively, B was a cause of C – the fact that the surgeon connected the finger at t_2 was a cause of the fact that x's finger was functional at t_3. Hence cause transitivity fails.[6]

The events in this example are specified in sentential formulation. But this is an incidental feature of the presentation of the example. The example holds just as well in nominalized form. Thus, the injury was a cause of the surgery, and the surgery was a cause of the functionality of the finger at t_3; but the injury was not a cause of the functionality of the finger at t_3. The force of the example is thus not dependent on whether narrow or broad event individuation is employed. The force of this example of the non-transitivity of cause also holds regardless of whether the world is assumed to be deterministic or indeterministic: cause is non-transitive in either case.

This example is, in particular, a counterexample to Lewis's counterfactual analysis of cause. Had the surgeon not connected x's finger (at t_2), the finger would not have been functional at t_3;[7] and had x's finger not been cut off (at t_1), the surgeon would not have connected it at t_2. Hence $\sim A > \sim B$ and $\sim B > \sim C$ are true. On Lewis's analysis, it should follow that A is a cause of C; but A is not a cause of C.[8]

Some purported counter-examples to cause transitivity do not, however, bring out the hard core of the failure of cause transitivity. Thus, consider a McDermott-type example,[9] in which a terrorist x placed a bomb, to be set off by x's pressing a button of a remote-control device. However, before x managed to press the button:

A – a dog bit x's right forefinger.

Unable to use his injured forefinger, and despite his pain:

B - x pressed the button with his left forefinger,

and consequently:

C – the bomb exploded.

Indeed, A is a cause of B and B of C, but A is not a cause of C.

This example, however, does not bring out the heart of the transitivity failure distinctive of being a cause since it trades on transitivity failure due to diagnonalizability. But failure of transitivity due to diagonalizability plagues causal relevance as well, not just the relation of being a cause;[10] and the core of the failure of cause transitivity is independent of diagonzalizability, as brought out in the above example. Thus, if A is a cause of B, and D is a cause of C, but B is not a cause of C nor A a cause of D, then A may well be a cause of $B.D$ and $B.D$ a cause of C even though A is not a cause of C. Such are cases of diagonalizability. McDermott's example brings out this failure of cause transitivity due to diagnoalizability. Consider:

B' – x pressed the button.

Then B is:

B – B' with his left finger.

The difference between B and B' expresses the *manner* in which the action of B was performed.

A, however, intuitively, is not a cause of B' (which is B with its manner specification deleted, though A is a cause of B), and the *manner* of B per se (over and above B') is not a cause of C.[11] So even though A is a cause of B and B of C, this is a case of diagonaliability, which is a pervasive source of failure of transitivity of the relations of being a cause as well as of causal relevance. Yet cause transitivity fails in a way that transcends diagonalizability, as shown in the counter-example in the beginning of this section: in it there is no diagonalizability, yet it is an instance of the failure of cause transitivity.

2. PROBABILISTIC ANALYSIS OF CAUSE

I now move to briefly sketch a probabilistic analysis of cause. The notion of *a cause*, as discussed and analyzed in this paper,[12] is considered strictly on the *token* level, namely, as relating particular event tokens. The corresponding generic relations are not discussed here. The probabilistic analysis of cause presented here is designed for a non-deterministic, probabilistic world, based on the notion of objective chances[13] (however, no assumption regarding the world's being probabilistic constrains the counterfactual analysis of cause presented below per se). The basic notion is that of the chance of an event C, given a prior history (or state) of the world $\boldsymbol{W_t}$,[14] symbolized as:

$$\mathbf{P}(\boldsymbol{C/W_t})$$

(for $t < t_C$, where $\boldsymbol{t_C}$ [15] is the time to which C pertains). The chance function P, as a probability function, allows for conditionalization. Thus, the basic locution we will use is the chance of C given the (earlier) event A and its prior history, that is:

$$P(C/A.W_A),$$

namely, the probability of C given A and $\boldsymbol{W_A}$, the latter being the history of the world up to A, i.e., up to t_A.[16] I will be concerned with causal relations between actual events, and thus assume that A and C are actual.[17] Since the account is probabilistic, I will employ narrow (as opposed to broad) event individuation.[18]

The problem to be addressed, then, is the following: given the actual events A and C, where A is temporally prior to C,[19] are there conditions (formulated in terms of chances) that capture A's being a cause of C, and if there are, what are they? If analyses of cause in terms of chances, such as the one I propose below, are successful, and assuming that chances are indeed objective, the analyzed causal notions are objective too, since they are analyzed exclusively in terms of chances. In this case, one metaphysical conclusion is that the pertinent causal phenomena are objective as well. Another metaphysical conclusion is that these causal phenomena are, ultimately, complex chance phenomena. Thus, to that extent,[20] the causal structure of the world is not independent of its nomic structure (taken here to depend on, or even include, its chance structure).[21] Yet it is not something that transcends the factual features of the world together with its nomic structure – not a separate, independent realm: the causal phenomena so reducible are complex facets of the factual and nomic (including chancy) character of the world.

The notion of cause analyzed here is the notion of *a* cause, not that of *the* cause.[22] *The* cause is one cause, selected from many. A given event has, in general, many, many causes, pertaining to various times prior to it, including times long before it occurred. The selection of one such cause out of the class of all causes of the event in question as *the* cause is in general context-dependent and interest-relative.[23] Though the notions of *the cause* and *a cause* are often conflated,[24] they are distinct. Nor do I analyze here the notion of *causing* (which is also pragmatic).[25] The notion of a cause will be analyzed probabilistically, and thus, if such an analysis is successful, it will come out as not context-dependent or interest-relative, and thus as objective in this sense – as a direct result of the objective nature of chance, in contradistinction to the notion of the cause (and of causing). In considering examples in the context of the analysis proposed below, and particularly in employing the associated intuitions, it is imperative to keep in mind that the notion of *a* cause is being employed, not the notion of *the* cause (or causing).[26]

In the next section I briefly expound an account of when A is a cause of C (A, C actual).[27]

3. EX POST FACTO PROBABILITY INCREASE

The major motivating idea for a probabilistic analysis of cause was that the occurrence of a cause increases the probability of the effect.[28] What this idea comes down to depends on the probability notion in question.[29] In our case, where the notion of chance is being used, with the chance of C being relative to some world history W_t,[30] the most natural idea is to interpret probability increase as:

$$(1) \qquad \mathrm{P}(C/A.W_A) > \mathrm{P}(C/{\sim}A.W_A).$$

That is, given W_A (the world history up to A), the probability of C given A is higher than the probability of C given $\sim A$.

(1), however, is not an adequate analysis of A's being a cause of C. The reason is formal: (1) is a function of W_A, but *not at all* a function of $\boldsymbol{W_{A,C}}$, the intermediate history between A and C.[31] Thus, (1), taken as an analysis of A's being a cause of C, yields that given W_A, A and C, A is (or is not) a cause of C regardless of what else transpires between A and C. This is clearly false: whether or not A is a cause of C is very much a function of what happens between A and C. Examples can be constructed with fixed A, C and W_A such that one intermediate history (say, $W'_{A,C}$) intuitively yields that A is a cause of C while another intermediate history (say, $W''_{A,C}$) intuitively yields that A is *not* a cause of C (as I have argued elsewhere[32]). An adequate analysis of what it is for A to be a cause of C must therefore be a function of the intermediate history $W_{A,C}$.[33]

The question is how to transform (1) in a way that takes $W_{A,C}$ into account. This is far from straightforward. A natural candidate might be:

$$(1') \qquad \mathrm{P}(C/A.W_C) > \mathrm{P}(C/{\sim}A.W_C),$$

(i.e., (1) with W_C replacing W_A). But surely (1′) will not do, because in general $\sim A$ is inconsistent with W_C: since A is actual, A is in W_C.[34] Attempts to tinker with this inconsistency are not promising. They are akin to the corresponding problem encountered in the analysis of counterfactuals: just removing $\sim A$ from the condition on the right-hand side will not do, since in general A is already there, as it is in W_C; and in general there need not be a unique (or even minimal and unique) set of statements in W_C the removal of which would lead to compatibility[35] with $\sim A$.[36] The challenge for the analysis here is how to make sense of probability increase of C by A in terms of chances in a way that takes the intermediate history into account in an appropriate manner.[37]

So we will call condition (1), which is *not* a function of the intermediate history, **ab initio probability increase** (in short ***aipi***): though this is the natural condition to start with, it will thus not do for an analysis of cause. In seeking an analysis of the notion of probability increase that takes into account the intermediate history, we must look for an **ex post facto** notion of probability increase, rather than an ab initio notion. An ex post facto probability increase is one in which the intermediate history is taken into account: a sort of *hindsight* probability increase, from a bird's-eye view,[38] with the intermediate history unfolded. While the notion of token cause is indeed rooted in probability increase, we must be specific about which kind. So, first:

THESIS 1. If A is a cause of C, then there is an ex post facto probability increase of C by A.

The task now is to spell out the requisite notion of ex post facto probability increase in terms of chances.[39]

In attempting to analyze the requisite ex post facto probability increase, my proposal is first to consider particular intermediate events. To begin with, suppose the ab initio probability increase condition (1) obtains. (A similar strategy will be pursued when (1) does not obtain, as we will see below.) (1) by itself merely indicates that A yields the requisite ex post facto probability increase of C.[40] However, since (1) is not a function of $W_{A,C}$, this indication is not conclusive. The question then is: Is there an intermediate event E that *reverses* the inequality in (1)? In other words, is there an actual[41] intermediate (i.e., in $W_{A,C}$) event E such that, if taken into account in the condition on both sides, that is, held fixed, yields '$<$' (instead of '$>$') in the thus-modified (1)? That is: is there such an E for which:

$$(2) \qquad P(C/A.E.W_A) < P(C/\sim A.E.W_A)$$

obtains?

If there is such an E, then the fact that (1) indicates an ex post facto probability increase is neutralized in view of E, with (2) indicating[42] instead an ex post facto probability *decrease*, and thus that A is *not* a cause of C. We may then say that *if* there is no such E, the fact that (1) indicates the requisite ex post facto probability increase is conclusive, and A *is* a cause of C. In this case we have thus already fulfilled the requirement of taking the intermediate history into account, since we quantified over the events that comprise it.[43] We have therefore already arrived at a sufficient condition for the suitable ex post facto probability increase, and thus for being a cause, as follows:

THESIS 2. If (1) obtains, and (2) does not obtain for any E, then the requisite ex post facto probability increase obtains, and thus A is a cause of C.

If, given E, A yields a probability decrease of C, as in (2), we will call such an E a **decreaser**[44] (for A and C).[45]

In analogy to the notion of an ab initio probability increase, (3) can be used to define **ab initio probability decrease** (in short: ***aipd***):

(3) $\mathrm{P}(C/A.W_C) < \mathrm{P}(C/{\sim}A.W_C)$.

Namely, *aipd* is a condition according to which the probability of C given A is lower (given W_A) than given $\sim A$.

In a case of ab initio probability decrease, there may also be an intermediate event E' that, once taken into account in the probability condition on both sides, reverses the direction of the probabilistic inequality, yielding a probability increase; that is:

(4) $\mathrm{P}(C/A.E'.W_A) > \mathrm{P}(C/{\sim}A.E'.W_A)$.

In analogy to the notion of a decreaser, an event such as E' will be called an **increaser**. The notion of an increaser will play a central role in our analysis. An increaser, then, defined when ab initio probability decrease obtains, yields probability increase when held fixed. An event that is either an increaser or a decreaser will be called a **reverser**.[46]

Reversers are thus defined when ab initio probability decrease or increase obtains. However, it is only natural to consider an intermediate event a decreaser or an increaser when (2) or (4) hold, respectively, even when (EP) holds; that is:

(**EP**) $\mathrm{P}(C/A.W_A) = \mathrm{P}(C/{\sim}A.W_A)$.

To illustrate the notion of an increaser and its role in the analysis of the notion of cause, consider the following:

EXAMPLE 2. The Comeback Team had been weak for quite a while, with poor chances of improving during the next season. Consequently, there were very favorable odds for betting that it would lose. Nevertheless, x bet \$$Y$ on its winning (A). Later, but before the beginning of the games, the Comeback Team unexpectedly received a major contribution, with which it was able to acquire a couple of first-rate players. Consequently, the team's performance was the best in the season (E), x won the bet, and C occurred: x improved her financial position.[47]

As of t_A, A yielded a decrease of the chance of C; thus *aipd*. But given E, A yielded an increase of the chance of C. Hence E is an increaser for A and C. And indeed, intuitively, A was surely a cause of C.

However, just as an increaser E_1 (given an ab initio probability decrease) reverses the indication of ex post facto probability decrease by the ab initio condition, and in turn indicates the requisite ex post facto probability increase, there may be another intermediate event E_2 that, when added to the condition on both sides, once again reverses the direction of the probabilistic inequality, yielding probability decrease:

(6) $\mathrm{P}(C/A.E_1.E_2.W_A) < \mathrm{P}(C/{\sim}A.E_1.E_2.WA)$.

We shall call such an E_2 a *decreaser for* E_1 (for A and C).[48] Thus, an increaser[49] does not conclusively indicate the requisite ex post facto probability increase. Analogously, neither does a decreaser conclusively indicate the corresponding ex post facto probability decrease, since a decreaser too may have an increaser *for it*. Increasers and decreasers may therefore be reversed by decreasers and increasers (for them), respectively, with neither conclusively establishing the requisite ex post facto probability increase or the corresponding ex post facto probability decrease.

Ex post facto probability increase is a function of the intermediate history. But a *mere* function, just *any* function of it, yielding probability increase, need not yield the requisite ex post facto probability increase: the latter must be *stable*, that is, not reversed by other intermediate events that are conditionalized upon. The presence of an increaser E with no decreaser (for it) constitutes a stable condition, since no intermediate event reverses the probability increase yielded through E, and thus constitutes a *conclusive indication* of the requisite ex post facto probability increase. Call such an event E a **strict** increaser. A *strict* increaser, then, is an increaser for which there is no decreaser.

Similarly, a decreaser for which there is no increaser conclusively yields the corresponding ex post facto probability decrease: the probability decrease it yields is again stable since the direction of the probabilistic inequality is not reversed by conditionalizing on other intermediate events.

Call such a decreaser a **strict** decreaser. That is, a *strict* decreaser is a decreaser for which there is no increaser.

The notion of a strict increaser is the central notion in our analysis of cause. I thus propose, as an analysis of the requisite notion of ex post facto probability increase, viz., that suitable for the notion of cause, the constraint that there be a strict increaser. Thus:

> A strict increaser for A and C yields the requisite ex post facto probability increase.

And accordingly:

> If there is a strict increaser for A and C, then A is a cause of C.

Yet a prerequisite for A's being a cause of C is that A be causally relevant to C. (A is causally irrelevant to C iff C is causally independent of A.) I therefore assume that A is causally relevant to C when analyzing the notion of A's being a cause of C.[50] Thus:

THESIS 3. If there is a strict increaser for A and C, then A is a cause of C (given causal relevance of A to C).

And accordingly:

THESIS 4. A strict increaser for A and C yields the requisite ex post facto probability increase (of C by A) (given causal relevance of A to C).

This notion of causal relevance requires an independent analysis.[51] Thus, in the Comeback Team example above (Example 2), E (the team's performance was the best in the season) was indeed not just an increaser but a strict increaser, and therefore A came out as a cause of C on our analysis, intuitively the correct result.

Above, we observed that in an ab initio probability increase case without a decreaser the requisite ex post facto probability increase obtains: the ab initio probability increase indicates the requisite ex post facto probability increase and thus A's being a cause of C, and this is not reversed by any intermediate event, and so is conclusive. This case is thus entirely analogous to the case of a strict increaser. For terminological unity, we will therefore call a case of ab initio probability increase a case with a **null** increaser.[52] If the null increaser is not reversed (i.e., if there is no decreaser), call the null increaser **strict**. Accordingly, Thesis 3 now applies to null strict increasers as well as to non-null ones (the case of null strict increasers was also accounted for by thesis 2 above).[53] Thus, we can conclude as follows:

THESIS 5. A is a cause of C iff there is a strict increaser (null or not) for A and C (given causal relevance of A to C).[54]

And accordingly:

THESIS 6. The requisite ex post facto probability increase of C by A obtains iff there is a strict increaser (null or not) for A and C (given causal relevance of A to C).

Analogously, a case with an ab initio probability decrease will be called a case with a **null** decreaser,[55] and if a null decreaser is unreversed (i.e., if there is no increaser), it will be considered a **strict** null decreaser. A strict null decreaser, like a non-null one, yields a case of ex post facto probability decrease.[56]

4. SOME POSITIVE CAUSAL RELEVANCE

The notion of mixed causal impact[57] is familiar: often an event A has both positive causal impact and negative causal impact on a later event C. Often, the positive causal impact is transmitted through one route, the negative causal impact through another. Suppose an agent performed an action A in order to bring it about that C (which in fact occurred). We can consider whether his action ended up being conducive or counter-productive to C. This involves assessing whether there was positive causal impact of A on C, negative causal impact, or both.

Consider an example:

EXAMPLE 3. A patient is in very poor shape, suffering from a liver problem and a lung problem. He is given a medicine (A) to help with his liver problem, but the medicine has deleterious side effects on his lung condition. Yet the patient's overall health improves (C). A had mixed causal impact on C: it had positive causal impact in improving his liver condition, and negative causal impact in aggravating his lung condition.

Mixed causal relevance consists, then, of some positive causal relevance and some negative causal relevance. My informal thesis regarding cause is:

THESIS 7. A is a cause of C iff A has some positive causal relevance to C.

The condition that there be strict increasers should thus be considered as an analysis of the notion of some positive causal relevance as well, thus:

THESIS 8. A has some positive causal relevance to C iff there is a strict increaser for A and C. (Again, causal relevance is presupposed.)

Theses 3 and 5 follow, of course, from Theses 7 and 8. The notion of some positive causal relevance thus receives a precise probabilistic analysis through Thesis 8.

It is highly plausible that some positive causal relevance is a necessary condition for being a cause; but whether it is a sufficient condition as well is far from obvious: it seems hardly possible to tell a priori[58] whether some positive causal relevance is in itself sufficient for being a cause or whether a stronger condition along this vein (say, overall positive causal relevance)[59] is required. An example might help realize that some positive causal impact is indeed sufficient for being a cause. Suppose that:

EXAMPLE 4. In a theater with an audience of 300, someone shouted "Fire!". Consequently:

> A – everyone in the audience scrambled to leave the theater.

The result was predictable:

> C - there was a tremendous commotion in the theater.

If x_i is an individual in the audience, consider: $A_i - x_i$ scrambled to leave the theater.

A is surely a cause of C, but I suggest that each A_i is also a cause of C: each A_i had some positive causal impact, however small, on C, thereby qualifying as a cause of C (one among many). (Obviously, the force of the example remains intact if one increases the audience size, e.g., by considering a stadium rather than a theater.)[60]

If an event A has some positive causal relevance to C but no negative causal relevance, we will say that A has **purely positive causal relevance** to C. Similarly, some negative causal relevance without positive causal relevance amounts to **purely negative causal relevance**. The notion of purely negative causal relevance plays an important role in the analysis of counterfactuals.[61] Also, causal relevance without positive causal relevance amounts to purely negative causal relevance, and causal relevance without negative causal relevance amounts to purely positive causal relevance. Many everyday cause-effect relations are not mixed: usually, A

is a cause of C when A has purely positive causal relevance to C. Thus, when x drops the chalk (A) and the chalk falls (C), A is a cause of C, and A has purely positive causal relevance to C. (In view of Thesis 7, it is sometimes helpful to consult our intuition regarding whether A has some positive causal relevance to C in order to buttress our judgment as to whether A is a cause of C.[62])

5. COUNTERFACTUALS

On the account of counterfactuals that I developed,[63] a central class of counterfactuals constitutes the *n-d* type,[64], i.e., counterfactuals $A > C$ such that the antecedent A is factual[65] and compatible with its prior history W_A.[66] I shall confine my discussion here to such counterfactuals, and more specifically, for the purpose of analyzing the notion of cause, to counterfactuals with a false antecedent and with standard temporal order, i.e., where the antecedent-event A is temporally prior[67] to the consequent-event C.[68]

Consider then such counterfactuals $A > C$ (A true). If we construe counterfactuals as governed by an inferential schema on which the counterfactual consequent is represented as inferable on the basis of certain *implicit premises*, the main question is what the schema's implicit premises are. On this conception, the counterfactual $\sim A, > C$ (A true) is true iff:

(7) $\{\sim A\} \cup \ldots \rightarrow C.$

The question is: what implicit premises should be substituted for the dots so that schema (7) is true iff the counterfactual is true?[69]

For counterfactuals of the *n-d* type, the history prior to A, $\boldsymbol{W_A}$, should be among the implicit premises; so do the laws of nature L. Thus, the inferential schema can be further fleshed out as:

(8) $\{\sim A\} \cup W_A \cup \ldots -L\rightarrow C,$

('$-L\rightarrow$' means inferability through the laws L.) The main problem is to specify what should be retained among the implicit premises from the actual history of the world during the (t_A, t_C)[70] interval – the *intermediate history* $\boldsymbol{W_{A,C}}$[71] – once the counterfactual hypothesis $\sim A$ is contemplated. Obviously, the truth value of the counterfactual depends, and strongly so, on what actually transpires during this interval: different intermediate histories may yield different truth values for a given counterfactual with the same prior history. Hence analyses of counterfactuals such as that of Lewis and others[72] which are explicitly independent of $W_{A,C}$ will not do.[73]

To see this, consider:

EXAMPLE 5. Suppose x is standing near the upper portion of the elevator shaft of the Hancock building in Boston, which happens to be empty all the way down. x holds a large sheet of paper in his hand. In fact:

A – x does not drop the sheet of paper at t,

and consequently:

C – the sheet of paper does not fall to the ground level.

Consider now the counterfactual $\sim A > \sim C$ (had x dropped the sheet of paper at t, it would have fallen to the ground level). y is standing on the 5th floor holding a barrier that, once inserted, completely cuts off the upper portion of the elevator shaft from the lower portion. Now consider the counterfactual under two different courses of events in the interval between A and C. In course 1, the intermediate event B occurred:

B – y inserts the barrier across the shaft at t+dt.

(The occurrence of B is independent of what x does with his sheet of paper.) In course 2, y does not insert the barrier. In course 2, the counterfactual $\sim A > \sim C$ is true. But in course 1, $\sim A > \sim C$ is false, in view of B. Hence the truth value of the counterfactual clearly depends on the intermediate course of events.

We are thus back with the question which statements describing the intermediate history[74] should be retained as premises in the inferential schema. In general, of course, the entire intermediate history need not be compatible with $\sim A$. The obvious suggestion is that what should be retained from the intermediate history once the counterfactual hypothesis $\sim A$ is entertained is *all the true statements that would still have been true had* $\sim A$ *been the case.* Namely, what should be retained are the consequents of the true semifactuals $\sim A > E$, where E is in $W_{A,C}$. (A **semifactual** is a counterfactual with a true consequent.) We thus need to have a closer look at semifactuals of this sort.

They fall into three groups. The first are semifactuals $\sim A > E$ such that A is causally irrelevant to E[75] (in other words, such that E is causally independent of A). Call such semifactuals **irrel-semifactuals**. The important thing about such semifactuals is that A's being causally irrelevant to E is a sufficient condition for the semifactual $\sim A > E$'s being true. Hence all irrel-semifactuals are true. As an example of an irrel-semifactual, consider the following:

EXAMPLE 6. The American astronaut on board the Mir space-craft in 1998 was eager to return to earth during the week prior to the rendezvous of Mir and the shuttle. However, prior to that, I sneezed (at t). Had I not sneezed at t, that astronaut would still have been eager to return to earth. This semifactual is true, and the fact that I sneezed at t was causally irrelevant to the later fact that the astronaut was eager to return. This is thus an irrel-semifactual, and, as such, true. In Example 5 above, regarding the elevator shaft in the Hancock building, in course 1, in which B holds (B – y inserts the barrier across the elevator shaft at t+dt), the semifactual $\sim A > B$ (i.e., had x dropped the sheet of paper at t, y would still have inserted the barrier across the elevator shaft at t+dt) is an irrel-semifactual (since B was causally independent of A) and thus, as such, true. Consequents of irrel-semifactuals are to be retained among the implicit premises of schema (8) above.[76]

The second group is made up of semifactuals $\sim A > E$ such that A has purely negative causal relevance to E. *Purely negative causal relevance* amounts to some negative causal relevance with no positive causal relevance.[77] The important thing about such semifactuals is that they are also all true: A's having purely negative causal relevance to E is a sufficient condition for the semifactual $\sim A > E$'s being true.[78] Call such semifactuals **pn-semifactuals** (*pn* – for *p*urely *n*egative).[79] As an example of a pn-semifactual, consider the following:

EXAMPLE 7. x shot y (E); but z struggled to prevent x from shooting (A), unsuccessfully. Had z not struggled to prevent x from shooting, a fortiori x would have shot y. This semifactual is true. A, the fact that z struggled to prevent x from shooting, was purely negatively causally relevant to E – the fact that x shot y. $\sim A > E$ here is thus a pn-semifactual, and as such true. Consequents of pn-semifactuals likewise should also be retained among the implicit premises of schema (8).

The members $\sim A > E$ of the remaining set of semifactuals may be true or false. Their crucial feature is that their truth-conditional behavior is just like that of non-semifactual counterfactuals: their consequents do not belong to the implicit premises, and thus are not privileged in the way that the above two sets of semifactuals are. Let us therefore call them *counterfactual-type-semifactuals*; for short: **con-type-sem**s, abbreviated further as: **cts**. Thus, the consequents E of true *con-type-sem*s $\sim A > E$

are not to be retained in schema (8) above: their truth value is determined non-trivially through their *own* inferential schemata (the analogues of (8)) – namely, through the retention of W_A and of consequents of irrel-semifactuals and pn-semifactuals with consequents in $W_{A,E}$.[80]

The division of such true semifactuals $\sim A > E$ (with E in $W_{A,C}$) into these three categories is thus exhaustive and exclusive. We can therefore complete the above inferential schema for counterfactuals in the following way (A being true):

$\sim A > C$ is true iff

$$(9)\quad \{\sim A\}\cup W_A \cup \left\{\begin{array}{l}\text{the set of true statements } E \text{ in } W_{A,C}\\ \text{where } A \text{ is causally irrelevant or purely}\\ \text{negatively causally relevant to } E\end{array}\right\} -L\rightarrow C.$$

To illustrate the importance of the retention of consequents of irrel-semifactuals, consider the following:

EXAMPLE 8.[81]. Someone (x) contemplates selling certain stocks. After deliberating, x instructs his agent (at t_1) to sell (A).[82] At t_2 (a couple of days later) the stock market skyrockets (E). At t_3 (the following day) it would (under a plausible fleshing out of the case) be true to say: Had x not sold his stock, he would have been able to retire. But for this intuitively true counterfactual to come out true on the above inferential schema (8), the event E – the stock market skyrockets at t_2 – must be retained among the true statements from which the counterfactual consequent is to be inferable (through the laws – see schema (8)). And indeed, the actual A is causally irrelevant to E. Just A and its prior history W_A do not suffice to yield the consequent $\sim C$ (C – x is not able to retire) through the laws.[83] A different intermediate history $W_{A,C}$ that does not include events such as E would not uphold this counterfactual. This example thus also establishes that the truth of $\sim A > \sim C$ hinges on whatever transpired between A and C – in particular, on events such as E. Hence the truth-conditions of $\sim A > \sim C$ must be a function of $W_{A,C}$. An analysis which is not a function of $W_{A,C}$ therefore will not do (e.g., Lewis's *Analysis 1*, which is not a function of the intermediate history $W_{A,C}$).[84]

To illustrate the importance of the retention of consequents of pn-semifactuals, consider the following example. An architect participated in a competition to design a museum. His design for the museum included elaborate ornamentation (A). The task of the board of the new museum was to choose a design from those submitted and the architect who submitted it as one package: the architect selected would be awarded a contract to build the museum according to the design he submitted.[85] The board, however, was reluctant to agree to the extra cost involved in elaborate ornamentation. The design competition was fierce. Yet, nevertheless, the board selected the above architect to build the museum (E). Thus, had the architect submitted his design without the ornamentation, a fortiori the board would have selected him to build the museum. The architect indeed built the museum as designed. But had his design for the museum not included the ornamentation, the architect would have then built the museum without the ornamentation. This counterfactual is indeed true. But for the counterfactual to come out true according to schema (8), information to the effect that E must be retained in the antecedent of schema (8). The fact that the design of the above architect for the museum included the ornamentation (A) had purely negative causal relevance to the fact that the board selected him to build the museum (E). Hence the consequents of pn-semifactuals must be retained for schema (8) to correctly capture the truth conditions of counterfactuals of this sort. Indeed, the semifactual 'had the design of the above architect for the museum not included the ornamentation, a fortiori the board would have selected him to build the museum' is a pn-semifactual, and as such true (the construction 'If . . . , then a fortiori _ _ _' is typical for pn-semifactuals.)

Lewis ignores the events in the (t_A, t_C) interval, preserving only the history prior to A and the laws.[86] As a result, his theory cannot in general handle cases that rely on the preservation of the above sets of statements describing intermediate events. Of course, for Lewis it would be conceptually difficult to accommodate such statements, since, in the first place, he does not seem to have a systematic conception of causal relevance.[87] But, more seriously, causal relevance and purely negative causal relevance are causal notions, yet Lewis is committed to reducing causes to counterfactuals, while the approach presented here calls for a reductive effort in the opposite direction.[88]

The notion of purely negative causal relevance has cognates. Consider a counterfactual $\sim A > C$ (A being actual). In schema (9), true factual statements E in (t_A, t_C) to which the negation of the antecedent, i.e., A, is purely negatively causally relevant are preserved. Not to confer purely negative causal relevance to E is either to be causally irrelevant to E, or else to have some positive causal relevance to E. The notion of some positive causal relevance therefore resurfaces in straightforward compounds of the notions of purely negative causal relevance and causal relevance. This notion is, on my account, of the highest importance since, according to Thesis 7, being a cause is having some positive causal impact.

Furthermore: on analysis (9) of counterfactuals, the statements to be preserved in the interval (t_A, t_C) are the true ones to which *A* is causally irrelevant or purely negatively causally relevant. But they are thus just the true ones to which *A* does not have some positive causal relevance. Hence they are just the ones of which *A* is not a cause. Analysis (9) above can thus be reformulated as an analysis of counterfactuals in terms of cause.

On the inferential schema as above, the counterfactual is true just in case the consequent is inferable from the antecedent and the implicit premises. Since we deal with an indeterministic world, inferability cannot be deductive. Otherwise, the counterfactual schema would not yield as true counterfactuals whose consequent is merely highly probable given the antecedent and the implicit premises. But such counterfactuals, when this probability is sufficiently high, comprise the bulk of counterfactuals we intuitively consider true. True counterfactuals whose consequents are entailed nomologically by the antecedent and the implicit premises constitute but a meager fraction of the counterfactuals we consider true, in particular causal counterfactuals.

Thus, the natural move is to consider inferability as involving sufficiently high objective probability or chance. Given the above analysis, the objective probability in question is the following: it's the chance of the consequent $\sim C$ given the antecedent $\sim A$ and the implicit premises; namely:

$$\mathrm{P}(\sim C/\sim A.i\text{-}p),$$

(where *i-p* is the implicit premises[89]). For a counterfactual $\sim A > \sim C$, the implicit premises are the ones specified above, namely, they consist of W_A and the statements in $W_{A,C}$, describing events that are causally independent of *A* or to which *A* is purely negatively causally relevant. The probability here is conditional chance, devoid of any epistemic elements. This probability is an objective feature of the counterfactual, to be called the **counterfactual probability** of $\sim A > \sim C$ (not to be confused with the probability of the counterfactual).[90] The natural move then is that the counterfactual is true just in case its counterfactual probability is high enough.[91] In this paper, however, I will proceed without relying on this aspect of the truth value of the counterfactual involving a high enough counterfactual probability.[92]

6. COUNTERFACTUALS AND CAUSES

I will now relate the above account of counterfactuals to the venerable counterfactual doctrine of cause, i.e., the theory that $\sim A > \sim C$ captures the notion of cause, possibly supplemented by closure under transitivity. Lewis upheld this conception for a deterministic world.[93] On this conception, a sufficient condition for A's being a cause of C is that $\sim A > \sim C$ is true; and a necessary and sufficient condition is that there be? a series of true counterfactuals $\sim A > \sim E_1, \sim E_1 > \sim E_2, \ldots, \sim E_n > \sim C$ ($n > 0$). On this analysis, cause is taken to be transitive; namely, if A is a cause of B and B of C, then A is a cause of C. I argued above (Section 1), however, that cause is not transitive, which, if true, undermines Lewis's analysis.

Yet if the above analysis of counterfactuals that I have proposed is correct, counterfactuals invoke causes. In view of that, to analyze cause in terms of counterfactuals is therefore, in general, circular. But I will ignore this problem here, and concentrate on the extensional adequacy of the counterfactual doctrine of cause.

To observe more closely that A might qualify as a cause of C even when the counterfactual condition $\sim A > \sim C$ is false (with apparently no chain of counterfactual dependence), and that consequently the counterfactual doctrine fails as a necessary condition,[94] consider an example:

EXAMPLE 9. Two individuals, #1 and #2, attacked and overcame a third, u, on a river bank, and then dragged him into the water. But #1 exercised more than sufficient strength for keeping the victim submerged, whereas #2 exercised by far less effort than what would be sufficient for that purpose;[95] and indeed:

C – u drowned.

Consider:

A – both attackers endeavored to keep u's head submerged.

and in particular:

A_i – attacker #i endeavored to keep u's head submerged.

Intuitively, it seems, each of the A_i's was a cause of C. In particular, A_2 (the fact that #2 endeavored to keep u's head submerged) was a cause of C (the fact that u drowned). Yet in this case, the counterfactual $\sim A_2 > \sim C$

(namely, if attacker #2 had not endeavored to keep *u*'s head submerged, *u* would not have drowned) is false.

In order to reinforce the point, consider a second variation in which #2 had enough strength to submerge the victim all by himself, even though in fact he exercised only a small portion of it, way below what was needed for him to submerge the victim all by himself; and further suppose that his disposition was such that, if necessary, he would have exercised the requisite force to submerge the victim. Still further, suppose that #1 did not exercise the requisite strength to submerge the victim all by herself, so that the contribution of #2 in fact made a real difference. This wrinkle does not alter what seems to be a robust intuition that A_1 was a cause of C. Yet in this case too, $\sim A_1 > \sim C$ is false. (Note that neither variation is a case of redundant causation.)[96]

In cases like this, where the relevant counterfactual is false, invoking cause transitivity seems of no avail: there does not seem to be an intermediate event B between A_2 and C (or A_1 and C, respectively, in the two variations) with counterfactual dependence of B on A_2 (or A_1) and of C on B. Hence such a counterfactual conception fails as an elucidation of the notion of cause. This is prevalent in cases with several contributing causes.[97] The counterfactual condition in such cases requires that, without the event A, the outcome C does not take place. Such a *sine-qua-non* condition (in cases where cause transitivity is of no avail) is too strong to capture the notion of cause:[98] the pertinent counterfactual (on the conception of a chain of counterfactual dependence, even assuming cause transitivity) is too strong to constitute a necessary condition for being a cause.

Yet in such cases the event in question, which is indeed a cause, clearly had *some* positive causal impact on the effect: in the second variation, A_1 (the fact that attacker #1 endeavored to keep *u*'s head submerged) had some positive causal impact on C (the fact that *u* drowned). The expectation that the notion of cause can be characterized by the counterfactual condition $\sim A > \sim C$ seems to indicate an expectation that being a cause involves positive causal impact that is indispensable or crucial for the occurrence of the effect. Despite this, in many cases causes merely have *some* positive causal impact on the effect.[99] Yet merely some positive causal impact does not suffice to ensure the truth of the counterfactual $\sim A > \sim C$, which undermines the counterfactual condition's being necessary for being a cause (even via a chain of counterfactual dependence with cause transitivity, were it to be endorsed, since often, in cases such as the above, there is no appropriate chain of counterfactual dependence). The falsehood of the counterfactual $\sim A > \sim C$ is entirely consistent with A's having some positive causal relevance to C, and thus with A's being a cause of C.[100]

Hence a condition sine-qua-non is not necessary for being a cause (even in an expanded form, via a chain of counterfactual dependence, and with assumed cause transitivity).

Note that in this example, in the second variation, $\sim A_1 > C$, which is true,[101] is a *con-type-sem*: as noted, A_1 is causally relevant to C and purely positively causally relevant to it (and accordingly it has no purely negative causal relevance to C).

Yet the intuitive appeal of the association of cause with the counterfactual condition $\sim A > \sim C$ is sufficiently strong that an explanation of the relationship between this counterfactual and the notion of some positive causal impact is called for, especially in conjunction with the account of counterfactuals I outlined above. The core of this intuitive connection is that the counterfactual $\sim A > \sim C$ is indeed a *sufficient* condition for A's being a cause of C.

Thus, suppose $\sim A > \sim C$ is true (A, C are actual). Then $\sim A > C$ is false (the truth of $\sim A > C$ is inconsistent with the truth of $\sim A > \sim C$). In particular, A is causally relevant to C (otherwise $\sim A > C$ would be true). Further, it is false that A has purely negative causal impact on C (which would be sufficient for the truth of $\sim A > C$). Since A is causally relevant to C, yet does not have purely negative causal impact on C, A has some positive casual impact on C. Hence A is a cause of C.

Thus, on the above account of counterfactuals, together with the thesis offered here, namely, that to be a cause is to have some positive causal impact, it follows, as a logical consequence, that if $\sim A > \sim C$ is true, then A is a cause of C, i.e., that the counterfactual conditional is indeed a sufficient condition for being a cause. This venerable intuition is indeed valid.[102] That this conclusion is a consequence of these two conceptions vindicates them both,[103] and in particular the latter. It also reveals the close affinity between the counterfactual construction and the notion of cause.

7. COUNTERFACTUAL TESTS FOR CAUSE

We have elaborated on the relationship between, on the one hand, the counterfactual $\sim A > \sim C$,[104] given the above analysis of counterfactuals, and, on the other, the notion of cause, taken conjointly with the thesis that cause amounts to some positive causal relevance. We saw that the truth of the counterfactual $\sim A > \sim C$ is sufficient for A's being a cause of C, a result that follows from the aforementioned accounts of counterfactuals and cause. This counterfactual test (namely, $\sim A > \sim C$'s being true) is widely used as a test for A's being a cause of C in philosophy as well as in law. However, we saw (Example 9) that various cases where $\sim A > C$

is true (and hence $\sim A > \sim C$ false) also yield A as a cause of C (where the resort to the ancestral is of no avail). Furthermore, we shall see below (Example 10)[105] cases where both the counterfactual $\sim A > \sim C$ and its corresponding semifactual $\sim A > C$ are false and yet A is a cause of C, which reflect (as we shall shortly see) the general feature that in *all* cases of the sort to which that example belongs A is a cause of C. From now on, we will focus on an indeterministic world (where overdetermination is of limited, if any, significance).

Consequently, the counterfactual test is unsatisfactory as a necessary and sufficient condition for being a cause (even when supplemented by the ancestral relation). But given the above accounts of counterfactuals and cause, the cause-counterfactual relation can be specified precisely, beyond $\sim A > \sim C$'s being a sufficient condition for A's being a cause of C. That is, it is indeed possible to provide a necessary and sufficient condition for being a cause in counterfactual terms.

In fact, more generally, all cases where the semifactual $\sim A > C$ is false (which cover the cases of $\sim A > \sim C$'s being true) are cases where A is a cause of C, regardless of the truth value of $\sim A > \sim C$. To see this, suppose the semifactual $\sim A > C$ is false. Hence $\sim A > C$ is neither an irrel-semifactual nor a pn-semifactual. Thus A is causally relevant to C, and A is not purely negatively causally relevant to C. So A has some positive causal relevance to C, and therefore A is a cause of C.

Thus, the fact that $\sim A > \sim C$'s being true is a sufficient condition for A's being a cause of C is a *special case* of the fact that the semifactual $\sim A > C$'s being false is a sufficient condition for A's being a cause of C (since the truth of $\sim A > \sim C$ implies the falsehood of $\sim A > C$, though not vice versa.)

Above (Example (9), the second variation) we saw a case of a *con-type-sem* $\sim A_1 > C$ where A_1 was a cause of C. The above reasoning establishes also that this is neither an isolated nor a merely prevalent case, but is true in general: for every *con-type-sem* $\sim A > C$, A is a cause of C. Thus, if $\sim A > C$ is a *con-type-sem*, then $\sim A > C$ is neither an irrel-semifactual nor a pn-semifactual. Hence A is causally relevant to C, and A is not purely negatively causally relevant to C. Hence A has some positive causal relevance to C, and thus A is a cause of C.

This reasoning holds regardless of whether the *con-type-sem* $\sim A > C$ is true or false. The class of false *con-type-sem*s overlaps with the class of false semifactuals. The complement in the class of semifactuals, for given A and C, of the class of cases in which $\sim A > C$ is a *con-type-sem*, consists of the cases where $\sim A > C$ is an irrel-semifactual or where $\sim A > C$ is a pn-semifactual. Hence, in all cases of semifactuals except the latter two, namely, in all cases of semifactuals that are *con-type-sem*s, A is a cause of C.

Examination of the law of counterfactual excluded middle makes clear that in many cases neither $\sim A > \sim C$ nor $\sim A > C$ is true.[106] Examination of counterfactuals in which the consequent is sufficiently underdetermined (in an indeterministic world) by the antecedent plus the implicit premises brings out once again the force of the claim that the strategy underlying the counterfactual doctrine of cause (even with the ancestral) is not promising. In particular, another variation on Example 9 can illustrate the case where both $\sim A > C$ and $\sim A > \sim C$ are false, in which case A is a cause of C.

EXAMPLE 10. Consider an indeterministic world, where the effort exerted by attacker #1 was well below what it would have taken for him alone to drown u, and attacker #2 used most of his strength in drowning u, and yet was disposed to do his best if acting alone, so that without #1, it was wide open whether #2 alone would have submerged u. $\sim A_1 > C$ is thus not true, and so is $\sim A_1 > \sim C$ (A_1 was: attacker #1 endeavored to keep u's head submerged; C was: u drowned.) Yet the intuition that A_1 is a cause of C remains robust (and resorting to the ancestral seems of no use here).

8. THE CHART

A is causally irrelevant to C iff C is causally independent of A iff C is *true independently* of A. That is, whether or not A happened, still C would have happened; hence, if $\sim A$, still C, if C is causally independent of A. Such counterfactuals can be read as: if $\sim A$, then still C (independently of A). Obviously, if A is causally irrelevant to C, A is not a cause of C.

If A is purely negatively causally relevant to C, the semifactual $\sim A > C$ is a pn-semifactual. In this case, there is no positive causal relevance whatsoever of A (vis-à-vis C) to be lost in switching to $\sim A$. Such semifactuals can be read as: if $\sim A$, a fortiori C: if A is purely negatively causally relevant to C, then indeed a fortiori C if $\sim A$ takes place rather than A. If A has purely negative causal relevance to C (which implies the truth of the semifactual $\sim A > C$), then, since A does not have any positive causal relevance to C, A is not a cause of C. If, on the other hand, a semifactual $\sim A > C$ is such that A is causally relevant to C and A does not have purely negative causal relevance to C, then A is a cause of C, since A has some positive causal relevance to C.

We can summarize these findings by providing a breakdown of the conditions for A's being a cause of C and for A's not being a cause of C in counterfactual terms of the above sort. This provides for a counterfactual analysis of cause, in the above terms, and also yields intuitive counterfactual tests for being a cause, since the truth-values of the counterfactual $\sim A > \sim C$ and the semifactual $\sim A > C$ are amenable to intuitive assessment (though fallibly so),[107] and the truth-values of irrel-semifactuals and pn-semifactuals can intuitively be determined by a fortiori-type semifactual formulations and independence-type semifactual formulations as well (see below). The category of true *con-type-sem*s cannot be directly determined intuitively, but can be determined indirectly by excluding the other groups through such intuitive counterfactual tests. This summary is presented in the following chart (where 'A c C' abbreviates 'A is a cause of C' and 'sem' abbreviates 'semifactual'):

The Chart:

$\sim A > \sim C$ true	$\sim A > \sim C$ false			
$\sim A > C$ false		$\sim A > C$ true		
		$\sim A, > C$ is:		
		true *con-type-sem*	irrel-sem	pn-sem
A c C		A c C	NO	NO

The reasoning behind this chart can be summarized as follows:

1. If $\sim A > \sim C$ is true, then $\sim A > C$ is false.
   If $\sim A > C$ is false, then there is causal relevance of A to C and A is not purely negatively causally relevant to C; hence A has some positive causal relevance to C, and hence A is a cause of C.
2. If $\sim A > C$ is true:
   a. If it is an irrel-semifactual, then A is causally irrelevant to C, and thus A is not a cause of C.
   b. If it is a pn-semifactual, then A is purely negatively causally relevant to C, and hence A is not a cause of C either.
   c. If it is a true *con-type-sem*, then A is causally relevant to C and A is not purely negatively causally relevant to C; hence A has some positive causal relevance to C, and hence A is a cause of C.

Put differently, for a semifactual $\sim A > C$:

1. If it is an irrel-semifactual, then A is not a cause of C.
2. If it is a pn-semifactual, then A is purely negatively causally relevant to C, and hence A is not a cause of C.
3. If it is neither, i.e., if it is either a true *con-type-sem* or not a true semifactual (the latter covers the case where $\sim A > \sim C$ is true), then A is causally relevant to C and A does not have purely negative causal

relevance to C, hence A has some positive causal relevance to C, and hence A is a cause of C.

The intuitive counterfactual test for being a cause is thus as follows:

1. If $\sim A > \sim C$ is true, then A is a cause of C.
2. If not, but if $\sim A > C$ is false, then, again, A is a cause of C.
3. Suppose $\sim A > C$ is true. Suppose further that if $\sim A$, then a fortiori C; then A is not a cause of C. Alternatively, suppose further that if $\sim A$ then still C (independently of A); then A is not a cause of C. Otherwise, A is a cause of C.

Another (perhaps shorter) way of expressing this counterfactual test of whether A is a cause of C is the following:
Check whether $\sim A > C$ is true.

1. If it is not, then A is a cause of C.
2. If it is, check whether 'if $\sim A$, a fortiori C' is true, or whether 'if $\sim A$, then still C (independently of A)' is true.
   a. If either one is true, A is *not* a cause of C.
   b. If neither one is true, then A is a cause of C ($\sim A > C$ is a *con-type-sem*).[108]

We can thus reformulate a precise counterfactual analysis of cause as follows:

> A is a cause of C iff:
> Either $\sim A > C$ is false, or else (if true), then it is false that if $\sim A$ then a fortiori C and it is false that if $\sim A$ then still C (independently of A).

Thus the notion of a cause can be analyzed in such counterfactual terms, and such an analysis yields a test for being a cause that can be decided using counterfactual intuitions.[109] (Yet, though useful in practice, this analysis is circular – see below).

9. HUME, LEWIS AND BEYOND

This analysis is very different from Hume's. For Hume, A is a cause of C iff $\sim A > \sim C$ is true. As we saw, this counterfactual is indeed a sufficient condition for A's being a cause of C, but is far from being a necessary condition: A may be a cause of C even when the counterfactual $\sim A > \sim C$ is false, either when both the counterfactual and the corresponding semifactual are false, or when the semifactual is true but is a *con-type-sem*. These two categories, that is, first, when the counterfactual and the corresponding semifactual are both false and, second, when the semifactual is a true *con-type-sem*, are categories where A is a cause of C and yet the counterfactual is false – categories not recognized by Hume.

The original three attackers example above (Example 9, second variation) features the true semifactual $\sim A_1 > C$ (had attacker #1 not endeavored to keep u's head submerged, u would still have drowned). A_1 was intuitively a cause of C, yet $\sim A_1 > C$ was a true *con-type-sem*. (In this example, attacker #2 had more than enough power to submerge u on his own and the determination to do so if necessary.) Example 10 illustrates the category where both $\sim A > C$ and $\sim A > \sim C$ are false, yet A is a cause of C. In it, it was wide open whether attacker #2 had the requisite power to submerge u by herself.[110] In this variation, $\sim A_1 > C$ is false, and so is $\sim A_1 > \sim C$. Yet intuitively A_1 was a cause of C.

The counterfactual analysis of cause presented here is also very different from Lewis's. For one thing, Lewis does not recognize features of particular facts in the intermediate history between A and C in his counterfactual analysis of cause,[111] and in particular, he does not recognize features involving causal irrelevance or purely negative causal relevance. Further, Lewis would be hard-pressed to recognize causal features in his analysis of counterfactuals, since such a move would risk rendering his analysis of cause in counterfactual terms circular. Lewis analyzed cause through the truth of the counterfactual plus transitivity, and thus presupposes that cause is transitive. However, as I argued above (Section 1), cause is not transitive.

In particular, compare the above analysis with a counterfactual-plus-transitivity account of cause (such as Lewis's ignoring now the non-validity of cause transitivity). The latter surely covers the slot in which $\sim A > \sim C$ is true (as yielding that A is a cause of C); but it covers neither the slot in which both $\sim A > C$ as well as $\sim A > \sim C$ are false, nor the true *con-type-sem* slot, whereas, as I argued, in both categories A is a cause of C. That in these two additional categories A is a cause of C I illustrated by Examples 9 (second version) and 10 above of the 2 attackers who drowned the victim. In these examples, $\sim A_1 > \sim C$ is false (A_1 – attacker #1 endeavored to keep u's head submerged; C – u drowned). In the first example (Example 9, second version), $\sim A_1 > C$ is true (since the other attacker was perfectly capable of submerging u without #1 and would have done so if acting alone) and is a *con-type-sem*. In this case A_1 is a contributory cause of C, and indeed intuitively, A_1 is a cause of C, despite the fact that $\sim A_1 > \sim C$ is false. In the latter example (Example 10), in which it was wide open whether #2, by himself, could submerge u,[112] $\sim A_1 > C$

is false as well, and yet intuitively A_1 is again a cause of C. But in both versions there does not seem to be any intermediate event that provides for stepwise counterfactual dependence connecting A_1 to C via transitivity. These examples therefore show that the counterfactual-dependence-plus-transitivity account holds neither in the case of the true *con-type-sem* slot nor in the case where both $\sim A > \sim C$ and $\sim A > C$ are false, since in neither case the example accommodates a rescue by transitivity.

It is well-known that the above account by Lewis faces difficulties regarding late preemption. It is thus worth noting that the above example is a case of contributory cause and not of late preemption.[113] The category of *con-type-sem*s covers contributory causes (such as Examples 9 and 10), cases of symmetric redundant causation (such as the firing squad),[114] and cases of asymmetric redundant causation such as early preemption (e.g., the standby assassin) and late preemption (see below).

We noted that Lewis's analysis of counterfactuals is deficient in being independent of the intermediate history between the antecedent and the consequent.[115] But, like counterfactuals, the relation of being a cause strongly depends on the intermediate history between the cause-candidate and the effect-candidate. Consider the following example:

EXAMPLE 11. x applied for a position in a certain department, where z was the chairman of the search committee. y, who was not a well-known philosopher, agreed to write a letter of reference on x's behalf; and indeed:

A – y wrote a favorable letter on x's behalf.

However, the following occurred:

C – x did not get the position.

However, y neglected, in course 1, to send the letter. So, in course 1, which we consider now, the fact that y wrote a favorable letter did not make any difference to the failure of x's application, and thus A was not, in this course, a cause of C.

However, in course 2, y's letter was sent, and E occurred (well before C):

E – z was told that y held him in low esteem.[116]

Consequently, having noticed y's letter in x's application, z vigorously opposed x's candidacy (B). In this course, A was indeed a cause of C, in particular if we assume, under indeterminism, that z's vote, as well as the overall vote, without y's letter, would have been an open question, and that z's vote was critical for the outcome:[117] the fact that y wrote a letter on x's behalf was a cause of the fact that x did not get the position.[118] Thus, whether A is or is not a cause of C is very much a function of what takes place in the intermediate course.

For Lewis, under indeterminism, the core sufficient condition for being a cause in terms of chance-consequent counterfactuals, holds in course 1 iff it holds in course 2, since it does not depend on the intermediate history.[119] In particular, so does the chance $\mathrm{P}(C/W_{\bar{A}})$, where $W_{\bar{A}}$ is the history up to the upper end of t_A: it is the same in both courses. On Lewis's analysis for the indeterministic case, for A to be a cause of C the actual chance $\mathrm{P}(C/W_{\bar{A}})$ must be much higher than the corresponding chances in the selected possible worlds, where these chances are $\mathrm{P}(C/W^i_{\sim\bar{A}})$ in the selected worlds W^i,[120] (which include W_A[121] and $\sim A$ and conform to the same laws as in the actual world).[122] But the selected worlds are the same whether we consider course 1 or course 2 in the actual world. Hence Lewis's analysis yields the same outcome regarding counterfactual dependence for courses 1 and 2, and there is no stepwise dependence on course 1. Regarding course 2, since in the example the decision of the search committee was an open matter given $\sim A$, there seems to be no stepwise dependence either. Hence on both courses Lewis's analysis yields the same outcome regarding whether A was a cause of C.[123] But intuitively, in course 1 A was not a cause of C, and in course 2 it was. This outcome reflects the fact that Lewis's analysis of cause in an indeterministic world is insensitive to the intermediate course.[124]

A frequent mistake made by jurists is to conclude from the truth of $\sim A > C$, and even from the falsehood of $\sim A > \sim C$, that A is not a cause of C. True, the truth of $\sim A > C$ implies that $\sim A > \sim C$ is false; but the fallacy in the inference is that redundancy cases are being ignored here (though they are not by philosophers) as well as cases of contributing courses. Thus, in many cases A is a cause of C even though $\sim A > C$ is true or just $\sim A > \sim C$ is false. To conclude from the falsehood of $\sim A > \sim C$ that A is not a cause of C is to fall into Hume's trap: it is to fail to recognize the two categories noted above – that in which $\sim A > \sim C$ and $\sim A > C$ are both false and that of true *con-type-sem*s – and to see that in both cases A *is* a cause of C.

Now that we have presented the counterfactual analysis of cause, displayed in the above chart, we can see that it rests on weaker assumptions than the ones we resorted to. Thus, for this counterfactual analysis all we need to require for the notion of cause is the thesis that being a cause amounts to there being some positive causal relevance. In particular, we did not rely on the probabilistic analysis of some positive causal relevance (or of cause): the derivation of the counterfactual analysis of cause does not depend on it. Further, we do not even need to assume the viability of a probabilistic analysis of cause. Similarly, regarding schema (9), the analysis of *n-d* counterfactuals, we do not need to assume probabilistic analyses of the notions that play a role in (9) of causal irrelevance or of purely negative causal relevance, nor do we need to assume that such analyses are feasible. Moreover, we do not even need to assume that the world is probabilistic: the analysis holds without this assumption. So the counterfactual analysis of cause offered here is, in these respects, general – transcending the conceptual framework with which I operated, namely, the one presupposing that the world is chancy.[125]

Furthermore: classical problems for the counterfactual analysis of cause center on cases of overdetermination,[126] or redundant causation, in a deterministic world. An analysis of cause couched in terms of chances (and accordingly, assuming an indeterministic world) that zeros in on and makes good use of the requisite probabilistic concepts, yields, I have argued, a satisfactory treatment of the problem posed by redundant causation, including preemption cases.[127] But the thrust of the above analysis, insofar as its main concepts are concerned, such as causal relevance, some positive causal relevance, and *con-type-sem*, whether fleshed out in terms of chances or conceived on a more general level, that is, apart from a probabilistic construal of these causal notions, is applicable to classical redundant causation cases as well.[128]

Thus, note that it was the basic category of *con-type-sem*s which captured the cases of A's being a cause of C, since the upshot of the above analysis was that *for A to be a cause of C is for* $\sim A > C$ *to be a con-type-sem* (true or false). Typically, in redundant causation cases, one event (at least), A, is a cause of C,[129] and thus $\sim A > C$ is a *con-type-sem* (and true); hence the specific importance of true *con-type-sem*s for overdetermination-type cases. In particular, in the case of the firing squad of two shooters, a paradigmatic case of symmetric redundant causation, where A_i is: shooter i shot, and C is: the convict died, both A_i's are causes of C, and the semifactuals $\sim A_i > C$ are indeed *con-type-sem*s by virtue of the A_i's being causally relevant yet not purely negatively causally relevant to C (and, moreover, true *con-type-sem*s).

Further, in cases of preemption (or potential over-determination, such as the early preemption case of the stand-by assassin or a late preemption case such as the shattered window and the two thrown rocks), the preempted cause B (which is, of course, not a cause) is causally irrelevant to C, and by virtue of that $\sim B > C$ is not a *con-type-sem* and B is not a cause of

C, whereas $\sim A > C$, where A is the preempting, and thus the real, cause, is a *con-type-sem* (and true). True *con-type-sem*s thus point to the actual cause(s) in both symmetric and asymmetric redundant causation cases and set the real cause apart from the merely potential cause in cases of the latter type. Thus, in the early preemption case of the Kennedy–Oswald example, with the standby assassin not shooting (since Oswald did), A (Oswald shot) is a cause of C (Kennedy was shot) and $\sim A > C$ is indeed a *con-type-sem*, since A is, of course, causally relevant to C and has some positive causal relevance to it, and, further, this *con-type-sem* is true.[130] Similarly, in a late preemption case where Suzy threw a rock (A) and so did Billy (B) (only a bit too late) and the window was shattered (C), $\sim A > C$ is a *con-type-sem* since A, which is a cause of C, has some positive causal relevance to C (and this *con-type-sem* is true), whereas $\sim B > C$ is not a *con-type-sem* since B (which is not a cause of C) is causally irrelevant to C.

NOTES

[1] See my *A Theory of Counterfactuals*, Hackett 1986 (currently published by Ridgeview; henceforth: **ATC**), ch. 2, and my 'Counterfactuals', *Erkenntnis* **36**, 1992, pp. 1–41.

[2] David Lewis, 'Causation', *Journal of Philosophy* **70**, 1973, pp. 556–567. Lewis's account is designed for deterministic worlds. Lewis retained these two tenets in his 1986 writings on causation (in his *Philosophical Papers*, Oxford University Press, Vol. II).

[3] See my 'Cause and Some Positive Causal Impact', *Philosophical Perspectives*; *Mind, Causation and World*, J. Tomberlin (ed.), Vol. 11, 1997 (henceforth: **C&SPCI**), pp. 401–432, esp. section 5.

[4] Lewis persisted in his commitment to cause transitivity also in his recent account; see his 'Causation as Influence', *Journal of Philosophy*, April 2000. (The present paper was processed by this journal before Lewis's latest paper was available.) For a criticism of Lewis's recent paper, see my 'Lewis' 'Causation as Influence', *Australasian Journal of Philosophy*, September 2001.

[5] See note 60 below for a further elucidation of this point in light of the analysis of cause below.

[6] One can produce a spectrum of examples of the same type regarding, e.g., some bodily organ or function, or some artifact or mechanism, that went awry and was restored. (I used this example in *C&SPCI*.)

[7] To reinforce this point, we can build it into the example that t_2 was the only time at which the surgeon was in a position to perform the surgery. Appropriate circumstantial constraints can, if deemed necessary, be similarly imposed to ensure the truth of the $\sim B > \sim C$. Note further that the force of the example does not hinge on temporal fragility – it does not depend on the construal of the instances t_1, t_2 and t_3 as instant-like: temporal intervals will do too.

[8] That $\sim A > \sim C$ is false illustrates the well-known non-validity of counterfactual transitivity.

[9] Michael McDermott, 'Redundant Causation', *British Journal for the Philosophy of Science* **XL**, 1995.

10 See my ‘Transitivity and Preemption of Causal Impact’, *Philosophical Studies*, **64**, 1991, pp. 125–160.

11 To see how this claim follows from my account of causal relevance and of cause, see my ‘Cause: Time and Manner’ (forthcoming), section 10. Thus, briefly:

$$P(C/B.W_B) = P(C/B'.W_B)$$

($W_B = W'_B$). This equality brings out the lack of probabilistic impact of the ‘difference’ between B and B', which amounts to the manner of B. Consequently, the manner of B is probabilistically irrelevant, and thus causally irrelevant, to C, and thus is not a cause of C. Furthermore, even though:

$$P(B/B'.A.W_A) > P(B/B'.{\sim}A.W_A)$$

and consequently A is a cause of the manner of B, A is not a cause of B', since:

$$P(B'/A.W_A) < P(B'/{\sim}A.W_A).$$

Yet the manner of B is not a cause of C, since:

$$P(C/B.W_B) = P(C/B'.W_B).$$

For further details, and more generally for a more extensive presentation of manner causation and time causation, see ‘Cause: Time and Manner’. For details of my account of probabilistic relevance, see my ‘Causal Relevance’ in *New Studies in Exact Philosophy: Logic, Mathematics and Science*, Vol. II., Bryson Brown (ed.), Hermes Scientific Pub. Co., 2000, or, for a short version, my ‘Causation: Probabilistic and Counterfactual Analyses’, in J. Collins, N. Hall, and L. Paul (eds.), MIT University Press, 2001, sections 3–6.

12 And similarly the notion of some positive causal impact; see below. I also assume factual realism. For a chancy universe, I take it, nomological realism boils down to objective chances not supervenient on the facts.

13 The notion of chance, as I use it, is not too different from Lewis’s. I used such a notion of chance for a probabilistic analysis of token causal notions in my Ph.D. dissertation (1975, University of Pittsburgh, microfilm). For more on this notion of chance, see my ‘Counterfactuals and Causal Relevance’, *Pacific Philosophical Quarterly* **72**, 1991, pp. 314–337, section 4, and my *ATC*, ch. 4, sections I and VI. For related conceptions, see K. Popper, *The Logic of Scientific Discovery*, Science Editions, 1961; I. Hacking, *The Logic of Statistical Inference*, Cambridge University Press, 1965; H. Mellor, *The Matter of Chance*, Cambridge University Press, 1971; B. Skyrms, *Causal Necessity*, Yale University Press, 1980; and especially D. Lewis, ‘A Subjective Guide to Objective Chance’, in *Ifs*, W. Harper, R. Stalnaker and G. Pearce (eds.), Reidel, 1981.

14 W_t is the world history up to t (or an appropriate world-state immediately preceding t). In a Markovian world it makes no difference which one is conditionalized on. I prefer the former due to its advantages in assessing particular cases, in view of our epistemic limitations.

15 t_C is the *reference time* of C: the time interval within which the event in question – *the C-event* – occurred (according to the sentence C). I assume that A, C, etc. are both sentences and the events they specify. Sentences specifying events thus determine (explicitly or implicitly) reference times for them. I also assume a non-relativistic framework, for simplicity.

In his ‘Counterfactual Dependence and Time’s Arrow’ (*Nous* **13**, 1979, pp. 455–476) Lewis points out, as one of his two reasons for rejecting what he calls *Analysis 1*, that certain counterfactual antecedents do not have reference times. When it comes to counterfactuals of the *n-d* type, employed in this paper (see below, Section 5), if there is no reference time (whether explicitly specified or implied by the context), the counterfactual is under-specified, and need not have a semantic value. It may be construed as a counterfactual type, different tokens of which (for different reference times) may have different truth values. (The same applies also for counterfactuals of the *l-p* type, not discussed here – see *ATC*, ch. 9, and my ‘Counterfactual Ambiguities, True Premises and Knowledge’, *Synthese* **100**,1994, pp. 133–164).

[16] t_A and t_C are, in general, temporal intervals.

[17] I also assume here (as does Lewis) that $\sim A$ is compatible with W_A.

[18] Such as Kim’s, where the paradigm is that of an ordered triple of an object, property and time, appropriately extended. For more, see ‘Causal Relevance’, note I occasionally ignore rigor regarding the use of quotes, above and below, for simplicity.

[19] I thus assume the relation of temporal priority; more specifically, that the occurrence time of the *A*-event starts before that of the C-event. The account of causation presented here is not designed to carry out Reichenbach’s program of defining temporal order through an independently analyzed causal order.

[20] That is, insofar as the causal phenomena in question are amenable to such analysis.

[21] Since the probabilistic features are central in such analyses.

Consequently, the causal structure is not analyzable exclusively in terms of factual features (‘is a cause’ is not reducible to ‘is’, to paraphrase Hume) if the chance structure does not supervene on the factual structure (as it need not, I hold, unlike Lewis, especially if the world is infinite).

[22] Although the plural forms of ‘a cause’ and ‘the cause’ – namely, ‘the causes’ (of certain events, two or more) – are grammatically the same, which makes it all too easy to conflate them.

[23] For example, in contexts where assignment of responsibility or liability is involved, one would typically limit the set of candidates for being *the* cause to human actions (so long as there are causes that are human actions), and in particular, exclude events prior to the births of the individuals involved, even though many such events qualify as causes (in the sense of *a* cause).

[24] See H. L. A. Hart and A. M. Honore, *Causation in the Law*, Oxford University Press, 1959.

[25] See H. Putnam, ‘Why There Isn’t a Ready-Made World’, in his *Realism and Reason*, *Philosophical Papers*, Vol. 3, Cambridge University Press, 1983, p. 214.

[26] For analyses of the notions of the cause and causing, see my ‘Cause: Time and Manner’.

[27] For comparison to other accounts, such as Lewis’s and Salmon’s, see *C&SPCI*, sections 1–4 and 15–17.

[28] The tradition of probabilistic causation has dealt largely with generic causes, whereas our concern here is token causes.

[29] The bulk of the treatments of probabilistic causation resorted to the relative frequency interpretation of probability.

[30] Or, if one wishes, to some world-state immediately prior to t_A.

[31] $W_{A,C}$ is the history from the beginning of t_A to the end of t_C.

[32] See my ‘Counterfactuals and Causal Relevance’, which makes this point for the case of causal relevance. But since cause implies causal relevance, the argument applies here as well. See also in *C&SPCI* the dictator example in Section 8, for illustration of this point for

the notion of cause. See also Example 11 below (Section 7), regarding the reference letter, illustrating that being a cause is a function of the intermediate history. A strong dependence on the intermediate history is also a feature of counterfactuals: see below, Section 5, esp. Example 5, and Section 7.

[33] Lewis's chance-consequent counterfactual analysis of cause is vulnerable to this criticism, since it is independent of almost all of the intermediate history (more specifically, the entire intermediate history after t_A) at least insofar as the sufficiency condition is concerned; see my discussion of his analysis in *C&SPCI*, sections 2, 15 and 16. A similar critical point applies to Mellor's analysis in his ...

[34] At least so long as $t_{\bar{A}}$ is earlier than the beginning of t_C.

[35] That is, consistency through the laws of nature.

[36] This line of thought regarding this type of problem originated with N. Goodman, *Fact, Fiction and Forecast*, Bobbs Merrill, 1955, ch. 1. See also *ATC*, chs. 1, 2 and 3.

[37] If one is inclined to possible-worlds analyses, one might consider closest possible worlds, or minimal changes, inspired by such approaches to counterfactuals. That is, one might suggest that, instead of a fixed history such as W_A or $W_{\bar{C}}$ (the history up to the upper end of t_C) or W_C to be conditionalized on, one consider the variety of $\sim A$-histories W^i 'closest' in some sense to the actual history. (Stalnaker and Lewis have both resorted to 'closest' possible-worlds in their analyses of counterfactuals.) But this does not seem to help, since the same problem recurs right away: If one opts for comparing $\mathrm{P}(C/W_{\bar{C}})$ with $\{\mathrm{P}(C/W^i_{\bar{C}}\}$ then the comparison is immediately undermined since C is in W_C, and thus $\mathrm{P}(C/W_{\bar{C}}) = 1$. Would one want to consider $\mathrm{P}(C/W_C)$ and compare it with the conditional chances $\mathrm{P}(C/W^i_C)$ in these selected worlds W^i? But then $W_{C,C}$ is not taken into account, and it may be a considerable (or sufficiently significant) portion of the history if t_C is long enough (or significant enough). But more importantly, greater difficulties arise since the probabilistic effects of A may be blocked, e.g., by effects of A that screen it off.

Apart from that, how does one extend the greater-than relation between two numbers to that between a number, say $\mathrm{P}(C/W_C)$, and a set of numbers? One way of doing so is to require that the given number $(\mathrm{P}(C/W_C)$ be higher (or even much higher) than any of the numbers in the set. (This method is similar to that of Lewis's in his analysis of cause by chance-consequent counterfactuals; see his *Philosophical Papers*, pp. 176–177.)

Such possible-worlds analyses, as sketched above, are designed not to be vulnerable to the fault found with (1) as they are indeed a function of $W_{\bar{C}}$. But are they objective? Is there an objective sense of 'the closest possible-worlds to the actual one up to $t_{\bar{C}}$'? Stalnaker, who adopted a closest possible-worlds analysis for counterfactuals, acknowledges that the choice of possible-worlds thereby specified is context-dependent. Lewis, too, opts for a context-dependent notion of similarity for the analysis of counterfactuals. One objective choice would be to include *all* possible-worlds that include W_A in which $\sim A$ is true (with the same laws as in the actual world). This was Lewis's choice for counterfactuals in an indeterministic world (ibid., pp. 180-181). Intuitively, however, this does not do justice to the notion of 'closest', since some such worlds seem more similar to the actual world than others, depending on features of $W_{A,C}$. But if 'closest' is interpreted as 'most similar', then one must ask: in what respect? Resorting to respects revives the context-dependent character of the similarity notion, since the weights or priorities associated with them seem context-dependent, again undermining the prospect of an objective notion of probability-increase, and hence that of an objective notion of cause. (Lewis himself does not offer criteria for selecting features of particular facts: in his analysis of counterfactuals, he avoids committing himself to the conception that similarity of particular facts makes a difference to overall similarity in the sense pertinent to counterfactuals (over and above perfect match – see his "Counterfactual Dependence and Time's Arrow", reprinted in his *Philosophical Papers*, Vol. II, p. 48), and he discards, this conception altogether in his chance-consequent counterfactual analysis of cause (*Philosophical Papers*, Vol. II, pp. 176-177, 180–181).

In *ATC* I argued that the resort to closest $\sim A$ possible-worlds (in the sense of involving minimal changes from the actual world) is incompatible with the analysis of counterfactuals I offered there. More precisely, I proved, on the basis of very weak assumptions, that the set of possible-world descriptions suitable for the core analysis of counterfactuals I offered there (the analysis of *n-d* counterfactuals; that is, the set of possible-world descriptions such that the counterfactual is true iff the consequent is true in all of them) is not represented by its subset consisting of those closest to the actual world (in the sense that this subset is not suitable for the analysis of counterfactuals if the former set of possible-world descriptions is); see *ATC*, ch. 7, sec. III, 6, pp. 172 ff., especially theorem 1, p. 179.

[38] To put an objective twist on the notion of hindsight, which is usually epistemically construed.

[39] There are in general different ways of articulating a notion of probability increase that satisfy the requirement of being a function of the intermediate history. I reserve the term 'the requisite ex post facto probability increase' for the kind that captures the notion of being a cause.

[40] The locution 'A increases the probability of C' is not intended to invoke any causal implication. Rather, it should be read here and elsewhere as an abbreviation of 'the probability of C given that A obtains is higher than given that $\sim A$ obtains' (and, of course, given W_A).

[41] Henceforth I only discuss actual intermediate events.

[42] Given (1), i.e., on the basis of (1) and (2) alone (given W_A).

[43] (2) itself is a function of the intermediate history since it is a function of the intermediate event E.

[44] That is: a decreaser, defined when ab initio probability increase obtains, is an intermediate event that yields probability decrease when taken into account on both sides.

[45] Note again: it is not that E *does* any decreasing; it is not implied that E causes anything in particular. Rather, given E, i.e., when E is taken into account in the condition on both sides, the probability of C is lower given A than given $\sim A$.

[46] I have not discussed here the issue of whether disjunctions can be causes. Similarly, I have not explicitly disallowed disjunctions as reversers. So long as no such limitation is imposed, one must protect against obvious trivializations. Thus, as is typically the case, when the chances at hand are not 1, $\sim A \vee C$ would constitute an increaser for any such A and C, and $A \vee C$ would constitute a decreaser for any such A and C, since adding the former would yield probability 1 to the left-hand side of (4) (without affecting the right-hand side) and adding the latter would yield probability 1 to the right-hand side of (2) (without affecting the left-hand side). (Such increasers and decreasers, if admissible, would also be strict.) A reverser is conceived here as an actual intermediate event that, given W_A, A and C, its occurrence or non-occurrence are not determined. So naturally $\sim A \vee C$ and $A \vee C$ do not fit this conception. Elsewhere, in discussing causal relevance neutralizers, it was obvious that, barring outright trivializability, a neutralizer must not yield C with probability 1 (given W_A) (see my "Aspects of Probabilistic Causation" (forthcoming), sections 5 and 6). At the absence of restrictions involving intermediate disjunctive events, such a restriction would be most to natural to impose here as well regarding reversers. Thus an increaser E would have to fulfill that $P(C/A.E.W_A) < 1$ (to disallow $\sim A \vee C$ as an increaser), and similarly a decreaser F would have to fulfill that $P(C/\sim A.F.W_A) < 1$

(to disallow $A \vee C$ as a decreaser); and similarly for decreasers and increasers for other reversers. (This constraint seems preferable to disallowing a reverser to pertain all the way to $t_{\bar{C}}$.)

47 Assume that x's financial state at t_A can be summarized by Z. Strictly speaking, C should be read as: x's financial position (at the end of the season) was better than Z.

48 As opposed to a decreaser (for A and C) *simpliciter*, which was defined above. The notion of an increaser *for* E_1 is defined analogously.

49 That is, simpliciter.

50 We must require causal relevance for the relation of being a cause since otherwise, even if there is a strict increaser for A and C, there might still be a causal relevance neutralizer, i.e., an event that neutralizes the would-be causal connection between A and C and renders A causally irrelevant to C. That is, the presence of a strict increaser is compatible with causal irrelevance. For an analysis of causal relevance and the notion of a neutralizer, see my "Causal Relevance", in *New Studies in Exact Philosophy: Logic, Mathematics and Science*, Vol. II., Bryson Brown (ed.), or, for a shorter version, my "Causation: Probabilistic and Counterfactual Analyses" (henceforth: **CPCA**), sections 3–6.

However, causal relevance and the presence of a strict increaser are necessary, but not sufficient, for being a cause. It must be further ascertained, when checking for causes, that there is no purely negative causal relevance despite the presence of a strict increaser, which can happen if the route of the would-be positive causal relevance (indicated by a strict increaser) is neutralized by a *positive relevance neutralizer*. This case can be analyzed with the notions used to analyze causal relevance (as in "CPCA" or in "Causal Relevance"), and I will not discuss it here; see my forthcoming "Partial Causal Neutralizers". I will ignore this complication below.

51 See "Causal Relevance" and "CPCA".

52 In a case of ab initio probability increase, an empty intermediate event may be considered a null increaser.

53 For Thesis 5 below to cover still further cases, the notion of a strict increaser must be further extended to include intermediate events E that yield a probability increase in a non-reversed way in an *aipi* case. See "C&SPCI", section 11.

54 Another principle governing the notion of cause is the Principle of *Cause Preservation under Conjunctive Expansion* (**CPCE**), namely: if A is a cause of C, then $A.D$ is also a cause of C, and A is also a cause of $G.C$ (as long as the temporal priority requirement is fulfilled; see note 19); or, in an extended form, *Cause Preservation under Informational Expansion* (**CPIE**), namely: if A is a cause of C, then A' is a cause of C' if A' and C' are informational expansions of A and C respectively (again, with the temporal priority requirement). (For a variant of this principle regarding causal relevance, see my "Transitivity and Preemption of Causal Impact", *Philosophical Studies* 1991, pp. 125–160. The account of causal relevance there has been replaced by the more general and much improved account of "Causal Relevance" or "CPCA").

55 Again, where empty intermediate events may be taken to be null decreasers.

56 For further wrinkles on the notion of strict increaser, see *C&SPCI*, section 11.

57 I use here the terms 'causal impact' and 'causal relevance' interchangeably. I do not distinguish here between A's having causal impact on C and A's being causally relevant to C. Surely, to have causal impact is to be causally relevant.

58 Namely, without examining a variety of cases. The notion of causal relevance is sometimes used in the literature on mental causation in the sense of positive causal relevance. I take this to be a misnomer, and in any case this is not the way I use the term: causal relevance can be positive as well as negative; see below.

59 For an analysis of this notion, see my "Overall Positive Causal Impact", *Canadian Journal of Philosophy* **24**, 1994, pp. 205–228.

60 In Example 1 for the non-transitivity of cause (Section 1), relatively small variations in C (C was: x's finger was functional at t_3) may avert A's not being a cause of the so-modified C; e.g., if C is modified into C': x's finger was almost functional at t_3; or into C'': x's finger was highly functional at t_3 (assume that prior to t_1 it was just functional). On a natural construal of the example, A (x's finger was cut off at t_1) had some positive causal impact on C' and on C'', and thus A was indeed a cause of C' and of C''. The fact that C is not stable vis-a-vis A's not being a cause of it, in the sense that A's not being a cause of C is quite sensitive to relatively small variations of C, does not tell against the force of the example: it merely brings out the feature that some positive causal impact can come in small increments.

61 See below, Section 5. In *ATC* I focused on $\sim A$ rather than A, and thus on purely positive rather than purely negative causal relevance.

62 That is, in certain cases we have clearer intuitions regarding some positive causal relevance than regarding cause.

63 See *ATC*, ch 2.

64 *n-d* – for *n*atural *d*ivergence; see *ATC*, ch. 2, sections III, IV.

65 That is, its indicative form is.

66 That is, the history of the world up to t_A. My discussion therefore focuses on the indeterministic case. See *ATC*, ch. 2, esp. V and VI, and "Counterfactuals".

67 In line with the Humean condition that causes are prior to their effects.

68 More specifically, regarding the notion of standard temporal order, $t_{\underline{A}} < t_{\underline{C}}$. $t_{\underline{A}}$ is the lower limit of the interval t_A (and similarly for $t_{\underline{C}}$).

69 Another option is to replace the inferential schema for counterfactuals, along with truth values for counterfactuals, with high objective conditional probabilities, where the probability of the consequent is conditional on the antecedent and the factual implicit premises, that is, the prior history and the intermediate events in (t_A, t_C) to which A is causally irrelevant or purely positively causally relevant (see below). The importance of these intermediate events (see below) is undiminished by this change in the semantic value of counterfactuals from truth values to conditional objective chances. If we denote the implicit premises (or their conjunction, say) as '***i-p***', this conditional probability associated with the counterfactual $\sim A > C$ can be represented as: $P(\sim C/\sim A.i\text{-}p)$. I will not discuss this conception here.

70 The interval starting at the beginning of t_A and ending at the end of t_C.

71 $W_{A,C}$ is the world history from the beginning of t_A till the end of t_C.

72 In his "Counterfactual Dependence and Time's Arrow" (p. 39) Lewis considers what he calls *Analysis 1* (which he attributes to Jackson, Bennett, Bowie and Weiner) and which is an analysis independent of the intermediate history (though this is not why Lewis rejects it). See also my "Counterfactuals and Causal Relevance".

73 However, my analysis is designed for the indeterministic case. Lewis's analysis for the deterministic case takes into account only miracles (major or minor) and perfect match (ibid., pp. 47–48). Yet also his counterfactual analysis applied to the indeterministic case is independent of the intermediate course, and thus the above critique applies; see his *Philosophical Papers*, Vol. II, pp. 176–177, 180–181. My critique of the above *Analysis 1* (note 72 above) applies first-and-foremost insofar as it is applied to the indeterministic case.

74 Which is, throughout this paper, actual.

[75] To avoid cumbersome formulations, I allow myself to invoke statements when, strictly speaking, it is the events they specify that are being discussed.
[76] An irrel-semifactual $\sim A > C$ can be expressed as: if $\sim A$, then still C (independently of A), which is equivalent to: C is true, independently of A.
[77] Or, put differently, causal relevance to E but no positive causal relevance to E whatsoever.
[78] See *ATC*, ch. 2, section VIII.
[79] In *ATC* I focused on the antecedent $\sim A$ and its relation to the consequent of the semifactual, rather than on the actual A, and therefore I talked about purely positive causal relevance and about pp-semifactuals rather than pn-semifactuals.
[80] Thus, there is no infinite regress; yet the extensional adequacy of schema (9) below is not compromised. For elaboration, see *ATC*, ch. 2, section XII, 2.
[81] Examples of this sort are examined in *ATC* (examples 3 and 8, ch. 2).
[82] I assume that the stock is sold in a private transaction (which remains confidential at least until t_3) to an individual who is not interested in further selling the stock and in fact does not sell (until t_3).
[83] But the doubling of x's stock investment would allow for his retirement. The stock market's skyrocketing is tantamount to more than doubling the value of the stocks across the board.
[84] "Counterfactual Dependence and Time's Arrow", p. 39.
[85] That is, the board was not empowered to propose changes in the design.
[86] For Lewis' account of counterfactuals in an indeterministic world; see his *Philosophical Papers*, Vol. II, pp. 176–177, 180–182.
[87] Peter Menzies ("Probabilistic Causation and Causal Processes: A Critique of Lewis", *Philosophy of Science* **56** (1989), 642–663) attempts to remedy this problem. Yet his dense counterfactual saturation is at best *sufficient* for causal relevance (since the counterfactual's being true implies causal relevance), but *not* the other way around. (See the categories of the analysis below, e.g., in the chart.)
[88] The approach I pursue is not circular. However, on my view, reducing causes to full-fledged counterfactuals (see Section 7 below) does not qualify as a non-circular reductive analysis, as counterfactuals hinge on cause-like conditions. (I here exclude chance-consequent counterfactuals, which need not be that way.)
[89] That is, their conjunction, or, if you will, the set comprising them.
[90] In various places I have noted the importance of this counterfactual probability and delineated the direction of moving to a position where the truth value of the counterfactual in an indeterminisitic world is determined by such counterfactual probability; e.g., in *ATC*, ch. 2, note 45, and my "Counterfactuals", note 2. An alternative to a commitment to counterfactuals' having truth values is to consider the counterfactual probabilities as the semantic markers of counterfactuals.
[91] Such a move has consequences regarding the logic of counterfactuals. Thus, accordingly, in view of the lottery paradox in a counterfactual form, it's clear that the counterfactual inference

$$\frac{\begin{array}{l} A > B \\ A > C \end{array}}{A > B \& C}$$

is not valid. Thus, by adding information to the consequent the counterfactual probability goes down. Further developments of these points must await another occasion.

[92] As I will indicate below (note 109), such a move does not affect the counterfactual analysis of cause developed in this paper.
[93] David Lewis, "Causation", *Journal of Philosophy* **70**, 1973, pp. 556–567. For the indeterministic case, he upheld an account based on chance-consequent counterfactuals, again closed under transitivity. For a critique of Lewis's position, see *C&SPCI*, sections 2, 15 and 16.
[94] See Lewis's related observation, in his *Philosophical Papers*, vol. II, p. 176.
[95] In a deterministic world. A suitable variant will do in an indeterministic world, e.g., by replacing this sentence with: But the strength excercised by #1 by himself yielded probability close to 1 of keeping the victim submerged, whereas the effort excercised by #2 alone would yield very low probability of keeping the victim submerged. (Note that the subjunctive formulation here should be construed in terms of conditional probabilities rather than counterfactuals.) The example, mutatis mutandis, holds in both cases.
[96] The intuitions that A_1 in this variation was a cause of C seems stronger than the intuition that A_2 in the former variation was a cause of C. This variation is suitable both for a deterministic as well as for an indeterministic case. Since the first variation in a deterministic world is not a case of redundant causation, it does not overlap with this well-known exception to the counterfactual conception of cause. Nor is the second variation a case of redundant causation, once we assume that #1 did not have enough strength to submerge the victim all by himself. This latter assumption does not alter the intuitive outcome that A_1 was a cause of C.
[97] Multiplicity of contributing causes is common and not an esoteric exception, and, contra Lewis, the intuition that in cases such as the above, especially the second variation, A_1 is a cause of C seems clear.
[98] As noted in Section 2, I do not discuss the notion of being *the* cause, which dominates many legal discussions of causation; see Hart and Honore, *Causation in the Law*.
[99] Not even *overall* positive causal impact: see my analysis of this notion in "Overall Positive Causal Impact".
[100] Assumed cause transitivity is quite effective in handling cases of early preemption.
[101] In particular, in a non-deterministic case (suitably adjusted), on which my interest is mostly focused.
[102] L. Lombard (*Philosophical Studies* **59**, 1990, pp. 195–211) argued that enablers do not cause the effect, yet the effect counterfactually depends on them, which he takes to show that the counterfactual account of cause is defective. Lombard may well be right that events of the sort he has in mind as enablers do not cause the effect; and this holds indeed over a larger array of causes. Here the distinction between the notions of a cause and of causing is crucial: a counterfactual analysis of causation, if it is to get off the ground, must first and foremost relate to the notion of a cause (which comports with Lewis's orientation), not directly to causing. An enabler may not cause the effect even when it is a cause of it.
[103] And in particular, the role played in them by the notion of purely negative causal impact, which has a close affinity with the notion of some positive causal impact.
[104] A, C true.
[105] In an indeterministic world.
[106] On Lewis's account and on my own. Stalnaker's conclusion is different. See R. Stalnaker, "A Defense of Conditional Excluded Middle", in *Ifs*, W. Harper, R. Stalnaker, G. Pearce (eds.), Reidel, 1980
[107] For instance, in case of ignorance about facts or probabilities.
[109] I assume that 'if $\sim A$, then still C (independently of A)' is a counterfactual formulation, expressing irrel-semifactuals.

The above counterfactual analysis of cause remains intact even when we consider the counterfactual as true just in case the counterfactual probability is high enough, since in either case, A is a cause of C just in case $\sim A > C$ is a *con-type-sem*. But which semifactuals qualify as *con-type-sem*s does not depend on whether or not we take a semifactual to be true just in case the couterfactual probability is 1 or merely high enough, since this depends on whether the appropriate causal irrelevance or purely negative causal relevance relations hold. (Irrel-semifactuals and pn-semifactuals are true under either option: the difference comes insofar as the truth-values of *con-type-sem* and of non-semifactual counterfactuals are concerned. But in all cases of *wh-type-sem*s $\sim A > C$, A is a cause of C, regardless of the truth-value of the semifactual.) If $\mathrm{P}(C/{\sim}A.i\text{-}p) < 1$ (A, C true), C is not preserved among the implicit premises, and thus $\sim A > C$ is a *con-type-sem*, and hence A is a cause of C. In particular, if $\sim A > \sim C$ is such that $\mathrm{P}(\sim C/{\sim}A.i\text{-}p)$ is high, then $\mathrm{P}(\sim C/{\sim}A.i\text{-}p)$ is low, and thus less than 1, and hence A is a cause of C. Likewise, if we abstain from assigning truth-values to counterfactuals in favor of holding the counterfactual probability as the semantic value, the conclusion that A is a cause of C just in case $\sim A > C$ is a *con-type-sem* remains intact.

110 Recall that I assume indeterminism here.

111 Lewis ascribes no importance to particular facts in his analysis of counterfactuals: under indeterminism he recognizes only perfect match (*Philosophical Papers*, Vol. II, pp. 180–182); under determinism, he acknowledges only miracles and perfect match ("Counterfactual Dependence and Time's Arrow", pp. 47–48). In his chance-consequent counterfactuals analysis of cause the intermediate course after t_A plays no role (*Philosophical Papers*, Vol. II, pp. 176–177).

112 And thus it is wide open whether C would or would not have happened if $\sim A_1$ had happened.

113 The unique aspect that differentiates late from early preemption regarding the counterfactual analysis of cause is the absence in the first of an intermediary that sustains step-wise counterfactual dependence. In the above cases of contributory causes there does not seem to be a suitable intermediary of that sort. The case of contributory causes is not a case of late preemption since in the former both contributory causes are genuine causes: neither cause preempts the other. (In cases of late preemption, when the courses of the two potential causes are independent of each other until very close to the occurrence of the effect, there arises the option of insisting on the fragility of the effect as a way out, albeit being an unsatisfactory one since insisting on the fragility of the effect is not a viable option when applied across the board, in particular in cases where the two causes-candidates are not independent of each other).

114 The case of symmetric redundant causation, such as the case of the firing squad, in a deterministic world, allows Lewis what seems to be at best a defensible position of claiming that neither potential cause by itself is a cause (though many, myself included, reject this intuition). But an analogous case in an indeterministic world allows for each shot to yield probability close to 1, but not quite 1, of the effect (the death of the convict). In such a case, a claim to the effect that neither candidate for being a cause is a genuine cause seems entirely implausible. (I am not aware that Lewis extends his above claim to the indeterministic case.) This case, like the case of late preemption, leaves no room for an intermediary that allows for a resort to cause transitivity, and thus does not accommodate Lewis' counterfactual treatment.

115 Again, in an indeterministic world. So would be counterfactual dependence (based on Lewis's conception of counterfactuals in an indeterministic world) if taken to be a sufficient condition for cause. Insofar as it is construed as a candidate for a sufficient condition for

being a cause. For Lewis's conception of counterfactuals in an indeterministic world, see *Philosophical Papers*, Vol. II, pp. 176–177, 180–181.
116 I do not assume that E was causally dependent on A.
117 I thus assume that in course 2 (though of course not in course 1) it ended up being a wide-open matter whether, without A, x would have gotten the position, and thus that $\sim A > C$ and $\sim A > \sim C$ are both false in course 2. $\sim A > C$ is of course true in course 1 (and consequently $\sim A > \sim C$ is false in course 1).
118 A is a cause of C even if we do not make the assumption of the previous note, that leads to $\sim A > C$ an $\sim A > \sim C$ both being false: Of course A is a cause of C if $\sim A > \sim C$ is true, but also if $\sim A > C$ is true, which is a *con-type-sem*.
119 More precisely, after $t_{\bar{A}}$.
120 See his *Philosophical Papers*, Vol. 2, pp. 176–177.
121 That is, the history just up to A.
122 Ibid, pp. 181–182.
123 Indeed, intuitively there seems to be neither counterfactual dependence nor stepwise dependence in either course.

A is a cause of C not only in course 2 of the above example, in which both $\sim A > \sim C$ and $\sim A > C$ are false, but also in a different variation in which $\sim A > \sim C$ is true – e.g., if x's record was comparatively so unimpressive that he would not have gotten the position even if $\sim A$. This is a case in which $\sim A > C$ is a *con-type-sem*.
124 At least after $t_{\bar{A}}$.
125 Some indeterministic assumption is required in order to allow for the compatibility of the negation of the counterfactual antecedent with its prior history. But since it was the basic category of *con-type-sem*s which captured the cases of A's being a cause of C, because the upshot of the above analysis was that for A to be a cause of C is for $\sim A > C$ to be a *con-type-sem* (see below), once the indeterministic ladder brought us to the above counterfactual account of cause, it may well be that the latter will do the job in the deterministic case as well. That is, the categories of causal irrelevance and purely negative causal relevance, in terms of which a *con-type-sem* was defined, may well be definable in a deterministic world as well.
126 Actual as well as potential, that is, including cases of pre-emption.
127 See "CPCA", sections 8 and 9.
128 Construed under determinism.
129 Regardless of whether the other potential cause-candidate is an actual cause as well.
130 For details, see "CPCA".

Philosophy Department
Hebrew University
Jerusalem
Israel