University of Alberta

Unified and Dis-unified Analyses of Causation

by

Jason D. Taylor ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of *Master of Arts*

Department of *Philosophy*

Edmonton, Alberta

Fall 2005

0-494-09072-3

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN:
Our file Notre reterence
ISBN:

for Mara

Abstract

Theories of causation can be classified into "global" and "local" accounts. These labels refer to the degree to which the theories make the causal relation between events sensitive or dependent on events 'outside' the immediate connection between cause and effect. While local theories (process theories) can analyse "preemption" scenarios, global theories (counterfactual theories) fail to give the correct analysis; and, while global accounts can naturally handle cases of causation by "prevention", local theories have difficulties. In the face of this division, two responses have recently been suggested: a *unified* account which tries to combine global and local elements into a hybrid analysis of the causal relation, and a *dis-unified* approach which proposes to distinguish *two* notions of causation so that each notion applies only to the appropriate class of cases. I argue that the dis-unified theory (Ned Hall) has serious problems with cases in which the different types of problems are combined into one causal scenario (as in "preemptive double prevention"). The unified account (Jonathan Schaffer), by contrast, can be modified so as to handle satisfactorily the whole spectrum of scenarios.

Table of Contents

1. Introduction

The late twentieth century theorists of causation stood arguing at a fork in the road: either the causal relation was best defined in terms of a local theory, or in terms of a global theory. The question was whether causation is a relation which was essentially definable in terms of the relationship between specific events; or whether the causal relation required the inclusion of 'other' factors in the world *besides* the specific relationship between the events in question. Local theories embodied the first idea, with the intuition that causation was a matter of processes between events (or whatever we take to be the causal relata). Global theories developed from the second idea; causation, the theorists thought, was a matter of more than the events in question – it was a matter of the events that go on elsewhere. Specifically, the intuition was that causation was a relationship of counterfactual dependence (or probability raising) between events. No matter the approach taken, the aim was to develop a theory such that the results which followed from it were in tune with our intuitions: when our intuitions tell us that Mara's action was the cause of some event, our theory had better agree (at least in most instances). For most cases, these theories did just that – and no solution was offered as to which road the causal relation would fare best on.

Local and global theories, however, were not without their own troubles as the theorist carried the relation down her preferred path. The theories ran into significant troubles, with respect to churning out conclusions which matched our intuitions, when it came to a certain set of situations. Now, every situation (not just the ones which are trouble for the first theories of causation) has a corresponding 'neuron' diagram (introduced in §2) which represents the structure of that causal situation. As we shall see presently, the structures revealed by the diagrams divide the cases into two sorts. The first sort, which we might call the local problems, are causal situations where the local theory gives the wrong result; the second sort, which we might call the global problems, are the causal situations where the global theory gives the wrong result. Each of these types has a paradigm example, upon which small variations can be rung such that the theorist (whether local or global), who is keen on fine tuning his theory so that it can provide a solution, will never be able to rest. The persistence of these problems – which have plagued both paths – has lead to general agreement that this fork in the road – local or global – is the wrong decision to be making.

Theorists have turned away from this fork, however, only to be confronted with another. Down the first road of the new fork is the prospect that there is one *unified* definition of causation which is able to churn out answers that correspond to our intuitions for all the cases using only one definition – achieved by combining, somehow, a process theory with a global theory. Down the other path, however, is not the prospect of a unified account, but rather a *dis-unified* account, which uses more than one definition to define the causal relation in order to achieve success.

There are compelling reasons to accept, and to reject, either of these paths; and, the issue, thus far, is stalled. I believe, however, that a close examination of the extant unified and dis-unified approaches will strongly suggest traveling down the unified-path. That is, any attempt at the development of a causal theory will be the most fruitful if it tracks the unified theorist's route, instead of the dis-unified theorist's. This conclusion is due to serious, crippling objections against the latter's approach. Further, I believe that a close consideration of the options, supplemented by the visual aids – the structural diagrams – will suggest how the unified theory can best be developed.

I start with some background information, sketching both David Lewis' influential counterfactual theory (§2.1) and some of the competing forms of process theories (§2.2). After showing where these theories fall short, I give a brief account of the second fork in the road (§2.3) before proceeding on to the leading accounts of each option. I discuss Ned Hall's dis-unified account first (§3.1), highlighting some of its major problems (§3.2). I then (§4.1) move on to Jonathan Schaffer's unified account showing where it, too, falls short (§4.2). In the final section, (§5), I offer a modification of Schaffer's account, and demonstrate how this account gives the intuitively correct solution for all known troublesome cases. In the same section I will suggest not only that my account is an attempt at incorporating stability into a counterfactual account (or counterfactual dependence into a stable – process – account), but also that, should my account fail in some way, any future theory of causation would best exploit a form of stability in order to be successful.

(A caveat: the literature on causation with which I deal is violent. It contains shootings, bombings, destruction, and death. I have made no attempt to alter the examples to friendlier, less bloody examples. While I leave uncharitable interpretations to the reader, I can offer one charitable interpretation for the prevalence of violence: in these examples our intuitions are quite strong; we are commonly saying that a certain person is to blame for some horrible action because she caused it to happen, and the abundance of these cases might serve to strengthen our intuitions. Now, I do not wish to go as far in my investigation as to explore the notion of responsibility, but I will milk our strong intuitions about these cases to guide my inquiry into a definition of causation. The reader is welcome, however, to rephrase each situation I discuss in terms of 'violent' free examples – such as by construing everything in terms of one neuron firing, causing another to fire, etc. – if she wishes. Personally, I do not find such abstract musings helpful or interesting.)

2. Background

2.1. Lewis' Global Theory

i) Deterministic Cases

Perhaps the most influential account of the causal relation is the one put forth by David Lewis (1973), which was an attempt to develop an idea that was as yet, undeveloped and under-appreciated: counterfactual dependence. Lewis proposed that causal relations are best defined by counterfactuals such that C is a cause of E if, and only if, had C not occurred, then E would not have occurred. Lewis symbolizes the counterfactual relation, $C \Box\!\!\rightarrow E$. (The current account is meant to cover cases of determinism. By "determinism," I follow Lewis by noting that I do not mean simply that every event has a cause, but rather that "the prevailing laws of nature are such that there do not exist two possible worlds which are exactly alike up to some time, which differ thereafter, and in which those laws are never violated" [1973, 196].)

E counterfactually depends on C when the following two counterfactuals are both true:

1) C□→E and,

2) ~C□→ ~E.

The Lewisian relation of causal dependence for actual events can then be defined in terms of (2), because (1) is automatically true. Hence,

> E causally depends on C if and only if C and E are actual, distinct events and ~C□→~E is true.

Determining the truth of the relevant counterfactual depends upon a comparison among possible worlds. When the material conditionals are true *in all the closest possible worlds*, we can say that C causes E. For Lewis this determination is made with respect to four criteria. One world is comparatively *less* similar to another if it differs by the presence of,

> *i*) widespread, diverse violations of laws (i.e, large miracles),
>
> *ii*) a lack of perfect match with respect to facts over a large spatio-temporal region,
>
> *iii*) small, localized violations in laws (i.e, small miracles), and
>
> *iv*) a lack of approximate match with respect to facts (Lewis 1973, 197).

To begin, as Lewis notes, the relation of similarity is a three-place relation between worlds: one world is farther *from the actual world* than another world in so far as it differs more drastically in the above four ways. Next, these four criteria have two formal constraints. First, the comparison of possible worlds is a weak ordering such that any two possible worlds can be compared for closeness to the actual world. Second, the actual world is understood to be closer in similarity to itself than any other world (i.e, the closest world to the actual world is the actual world itself.)

For the counterfactual $C \Box\!\!\rightarrow E$ there are, strictly speaking, two situations for possible worlds which we need to consider. The first world is a world where C holds (called a C-world) and where E also holds; the second is a C-world where E does *not* hold. If the C-world where E holds is closer to the actual world than the C-world where E does not hold, then the counter factual is true. If the C-world where E does not hold is closer to the actual world then the counterfactual is false. That is, we must find if "it takes less of a departure from actuality to make the consequent true along with the antecedent, than it does to make the antecedent true without the consequent" (Lewis 1973, 197). Similarly, for the counterfactual $\sim C \Box\!\!\rightarrow \sim E$: if the closest non-C-world is one where E also does not hold (as apposed to one where E does hold) then we can say that the counterfactual is true.

With all of this in mind, let us consider a simple causal relationship about which we have solid intuitions, apply these criteria, and see if Lewis' theory gives us the correct answer. Pretend that in the actual world, the president of the United States pushes the button to launch a powerful nuclear device and the world is destroyed.[1] Our cause (C) is the president's push of the button; our effect (E) is the destruction of the world. Now, to determine whether E causally depends on C we must ask whether 1) "if

[1] This example, originally given as a counter example to Lewis' theory of counterfactuals, is discussed by Lewis in (1979).

the president had pushed the button, the world would have been destroyed" is true; and whether 2) "if the president had not pushed the button, the world would not have been destroyed" is also true.

Because of the second constraint for the comparative similarity of worlds, i.e, that the actual world is the world which is closest to itself, (1) is automatically true. Thus, to determine whether E causally depends on C, we must evaluate the counterfactual, found in (2), $\sim C \Box\!\!\rightarrow \sim E$. We must ask, that is, if "if the president had not pushed the button, the world would not have been destroyed" is a true statement.

As an aid to evaluating counterfactuals, we might ascribe to (i) - (iv) values which would indicate the degree of difference (from the actual world) that a violating world has. Thus, e.g., a world that has one large miracle would get four points; a world that lacks a perfect match with respect to facts over large spatio-temporal regions might get three points; a world that has a small miracle might get two points; and finally, a world that lacks an approximate match with respect to the facts might get one point.[2]

Recall that what matters is if the closest non-C-world is one where E does not hold either (rather than one where E does hold), and then consider the counterfactual in question: if the president had not pushed the button, the world would not have been destroyed ($\sim C \Box\!\!\rightarrow \sim E$). Lewis states that we hold the antecedent fixed (as false) along

[2] Such a rigid designation of values will not do any work for solving the problem of effects (which I shall discuss shortly): Under these numbers, the cause will come to be counterfactually dependent on the effect. We could modify the ratings slightly to achieve the desired result – one which fits with Lewis' approach – if so desired. As this is not Lewis' approach *per se*, I will not try to modify the numbers to fit all situations; nor will I use the rankings beyond this example.

with the rest of the world, and observe the effects on the consequent. In other words, we compare two worlds with our own. The first world, W_1 is where C is false, and E is false. W_1 would have *(ii)* a lack of perfect match with respect to facts over a large region of space-time, given that for all times after the time of C, W_1 would differ from the actual world. To see this, recall that these criteria apply to world where determinism is true, and that under these stipulations, the differing event (the non-occurrence of C as opposed to its occurrence) would drastically change every event that followed it. A violation of *(ii)* counts as three points for W_1.

Given that we are only considering the possible worlds which are closest to the actual, W_1 will be one that is exactly similar with respect to facts in the actual world, up until the time of ~C; only after this time will it vary, (perhaps) greatly, because of determinism. If two worlds are exactly the same up to one point in time, and these worlds have the same laws (as do W_1 and the actual world, *ex hypothesi*) then they cannot differ in facts after that time without some (minor) violation of the laws of the world. As such, W_1 will also differ from the actual world by having *(iii)* a small miracle. This miracle, of course, is necessary to explain C's non-occurrence, and is worth two points for world W_1: our total is five.

Finally, W_1 will also have *(iv)* a lack of approximate match with respect to facts, raising W_1's total to six. There will not be any *(i)* large miracles, so our final sum will remain at a factor of six.

The second world, W_2, is where C is false and E is true. W_2 will be similar to the actual world up until the time of non-C. Afterwards, like W_1, a small miracle will have to occur in order that C not happen (two points). After non-C, W_2 will differ until the time of E, at which point E will actually occur (i.e, the world will be destroyed). Given that C did not occur (contrary to the actual world), and given determinism, what should have followed is non-E. But, because what does follow is E, thus, *another* small miracle must occur (two more points; W_2's total is now 4). After ~E occurs, W_2 and the actual world will be exactly the same given determinism (and barring further miracles). Accordingly, W_2 will lack a perfect match with respect to facts over a large space-time region (an additional three points), but will *not* lack an approximate match (given that the worlds only differ for the comparatively small time between the non-button-pushing event and world-destruction event. In fact, even if there is a large duration of time between when the time at C and the time at E, there will still be an approximate match with respect to facts, because we compare time lines over infinitely large intervals.). This brings the W_2's comparative difference up to seven.

As we can see, W_1 – lacking both a perfect and an approximate match in facts, and having one small miracle – is closer to the actual world than world W_2 – which has a lack of a perfect match in facts, and *two* small miracles. As such, we are entitled to say that the counterfactual ~C□→~E is true in the closest possible world to the actual, and hence that E counterfactually depends on C. And, because when both C□→E and ~C□→~E hold true, we can say that E causally depends on C. In other words, C is a cause of E because both the counterfactuals (C□→E and ~C□→ ~E) are true in the closest possible worlds.

These four criteria are not all that are required for interpreting Lewis type counterfactuals, however. According to Lewis it is also necessary to stipulate that the counterfactuals are of the non-backtracking type. A case of backtracking occurs when we can say that a *cause* depends counterfactually on its *effect.* This occurs in situations such that C could not have failed to cause E if C occurred; and, if E could not have been caused by any event other than C (Lewis 1973). Here, if E had been absent we know that C would not have occurred (because had C occurred, E would have): thus, *C counterfactually depends* on *E* ($E \Box\!\!\rightarrow C$); if this relationship holds, then E is a cause of C. This is the "problem of effects" (1973, 203), is a question of backtracking, and we must deal with it.

Lewis solution is to "flatly deny the counterfactuals that cause the trouble" (1973, 203), *even if* the relationship between C and E is such that C could not fail to cause E. We can say only that $C \Box\!\!\rightarrow E$, and not that $E \Box\!\!\rightarrow C$. To see Lewis' justification for this, consider the truth of the counterfactual $E \Box\!\!\rightarrow C$ at the closest possible worlds. That is, consider whether the world W*, where E is false and C is false is closer to the actual world than a world W**, where E is false and C is true. The desired result from such considerations, of course, is that W** be closer because this would mean that the counterfactual $\sim E \Box\!\!\rightarrow \sim C$ is false, and thus that the relation of counterfactual dependence does not hold. Though this conclusion – that $\sim E \Box\!\!\rightarrow \sim C$ is false – is not immediately obvious, a little thought will show that it is the correct one to make.[3] Take two worlds. These worlds are exactly the same as the actual world up until the time of

[3] Here the ranking system I used above does not work because I have not structured the values properly in order to accommodate for the *degrees* of similarity with respect to facts *over differing amounts of time.*

C, at which point W* (where C is false) will diverge, and W** (where C is true) will remain similar. At all subsequent times, W* will lack similarity to the actual world up to, including, and beyond the (later) time of E. However, it is only at E that W** will begin to diverge. Even though the divergence in both worlds W* and W** requires a small miracle, and both worlds have a lack of perfect match and approximate match with respect to facts over a large region of space-time, W** is still closer to the actual world than W* because it remains similar for a longer period of time. Indeed, even if the events C and E are temporally very, very close to one-another, W** would still be more similar to the actual world than W* by that fraction (of a fraction) of a second. This procedure results in what has come to be known as denying back-tracking counterfactuals, and is how the problem of effects is generally disposed of.

ii) Probabilistic Cases

Probabilistic theories, in general, say that some event is a cause of another if the occurrence of the first event increases the probability that the second will occur. Recall that global theories, as I have been presenting them, are concerned not with the particular connections between events for any situation, but with what goes on else where in the world. Probabilistic theories are true to this characterization; because what is relevant is that one event raise the probability of another, and because this probability can be determined only by measuring frequencies between the two events *across the whole world*, what goes on elsewhere is *vital* to the probabilistic theory. As such, it too is properly a 'global' theory.

Adapting Lewisian counterfactuals to the probabilistic case takes on many of the properties of the deterministic counterfactual theory, and many properties of the purely probabilistic theory. For instance, Lewis incorporates the idea that the cause raises the probability of its effect. And, even though the Lewisian counterfactuals change appearance under the guise of a probabilistic theory, the procedure with respect to possible worlds is (relatively) the same as in the deterministic case. The changes, when all is said and done, provide the definition that E depends causally on C, if and only if,

3) had C occurred, E would have had some chance X of occurring;

4) had C not occurred, E would have still had some chance Y of occurring; and,

5) X is, "by a large factor", greater than Y. (1986b, 177)

First, in the indeterministic case – as in the deterministic case – whether we can say that an event E causally depends on another event C still depends upon whether the counterfactuals given in (3) and (4) are true. This determination is still made by utilizing the Lewisian criteria (*i*) – (*iv*). But, these criteria for determining the closeness of worlds are to be understood such that they do not conflict with indeterminacy (e.g., the laws of the world must be indeterministic laws).

Second, the qualification "by a large factor" needs explaining. As Lewis says, it is not that *the difference* between X and Y must be large in order for C to count as a cause of E. Imagine that I plant a genuinely chancy bomb (it is set to read a radioactive decay device, say), but that the chance of the bomb actually exploding is quite small; around 0.1.[4] Let us say that the bomb, nevertheless, explodes. In this case, because X (0.1) is small to begin with, the *difference* between X and Y – Y being the chance that the bomb has of exploding when I do not hook it up to the decay device – could not have been large. If the chance of the bomb exploding, when it was not linked to this particular radioactive device was 0.01, then the difference would be a mere 0.09 – not large at all. However, my actions have increased the bomb's explosion by a *large factor* (i.e, a factor of ten). On these grounds, my actions ought to count as a cause of the bomb's explosion.

Third, it is important to remember that probabilities, as understood by Lewis, are "time dependent." In a string of events (not all of which are necessarily connected with every other event), relevant probabilities can, and likely will, change as time progresses. Thus, "the actual chance X of E is to be its chance immediately after C; and the counterfactual is to concern chance at that same time" (1986b, 177).[5] This chance, X, along with other chances in the world, supervenes on the history of the world up to that point in time. That is to say, a difference in facts in the world (as compared to the actual world) can affect the chance of a given later event in that same counterfactual world.

Finally, this revised counterfactual theory can be applied to the deterministic world. This is obvious: the chance X of E occurring when C occurs in a deterministic world, is just to be understood as 1; and, the chance Y that E had of occurring when C

4 This is Lewis' example (1986b, 176).

5 Small changes have been made in the quote to maintain consistency with my own paper.

does not occur in the same world, would be 0. It is a strength of Lewis' theory that it can be applied to both the indeterministic and deterministic situation.

Let us call this revised probabilistic account of counterfactuals (given in [3]-[5]), and the deterministic account given first, *global theories of causation*. The name invites the characterization that whether C is a cause of E depends on what is going on *elsewhere* in the world; the specific causal structure (of the events) in question is not the only thing that is relevant in our determination of C's causal character with respect to E.

A particular strength of a global theory – specifically Lewis' – is its ability to handle causal situations which involve omissions and non-occurrences of events. That is to say, not only can the theory replicate our intuitions about simple causal situations (like whether the pushing of the button caused the destruction of the earth), but it can also do the same for situations which involve troublesome omissions of events.

An omission is the non-occurrence of an event. Sometimes we like to say that omissions can be caused: the rain prevented the fire (i.e, the rain caused the fire not to occur); other times we like to say that an omission can cause some actually occurring event: the father's ill attention caused the child to be hit by a car (i.e, his omission of precautions caused the child to be hit by the car); and even other times we like to say that an omission prevented an event from occurring: because you did not push me, I was not hit by the car (i.e, the omission of the push-event by you, prevented the hit-by-a-car event).

Omissions are a peculiar type of non-actual event. The ontological status of an absence is troubling because it is not like that of the non-button-pushing event, or the non-world destruction event (from the example above) of a possible world. An absence is, in some sense, an actually non-occurring event that has a different status from other possible events. Lewis, however, sidesteps the issues concerning the ontological status of omissions by recognizing that for any event (or omission of an event) there is a proposition, which we may call a *proposition of occurrence*,[6] which corresponds to that event. For "any possible event E, there corresponds the proposition O(E) that holds in all and only those worlds where E occurs" (Lewis 1973, 199). Similarly, the proposition ~O(E), will hold in all and only those worlds where E does not occur. Then, "counterfactual dependence among events is simply counterfactual dependence among corresponding propositions" (*ibid.*). This procedure, strictly speaking, is to be used in all cases whether or not the events in question are omissions or not; however, I will continue to use the language of events (for ease of speech) so long as the reader bears this small modification in mind.

Even if we adopt this approach, we still have to say something concerning the abundance of causation by omission; if there is causation by omission, then there is lots of it. Consider some counterfactual situations: if, for instance, my hard drive had burned up, or if there had been nerve gas in the air while I was typing, or if my computer had been stolen, or if I had had a heart attack, then I would not have finished this essay. Thus, the failure of my hard drive to burn up, the omission of nerve gas in the air while I was typing, the failure of my computer to be stolen, and that I failed to

[6] This term is actually coined by Collins (2004, 109)

have a heart attack, are *all* causes of my finished essay. Indeed, we can even take the cases to an more absurd level: what if I had been kidnapped by aliens or turned into a gerbil by a disgruntled wizard?

Presumably there are lots of events which could have prevented my essay from being finished. These are what Lewis calls "overly disjunctive events" (1986b, 190). An event is overly disjunctive "to the extent that it could have occurred in various dissimilar ways" (*ibid.*), and must be avoided especially if we hold that there are events which are essentially specifiable as omissions. (Thus, when I say that Bonny omitted the precautions, there is an event 'the omission of precautions' which occurs no matter what Bonny does to omit the precautions, e.g., working, or reading.) Making the distinction between acceptable event and highly disjunctive events requires drawing the line somewhere; Lewis is unsure on where exactly to draw this line.[7]

One suggestion for understanding omissions in general is to treat them as causes, but *not* as essentially specifiable as such. So if Bonny omits the precautions because she was working, her working was "a genuine event; it is not objectionably disjunctive" (1986b, 192). This way, working, was a genuine event, and was the one way (of all the possible ways that she might have) that she omitted the precautions.

Nevertheless, this approach requires a special understanding of counterfactuals because, to suppose that Bonny did not work is not to suppose at the same time that the precautions were then made. As Lewis says, "it is one thing to suppose away an event *simpliciter,* another thing to suppose it away *qua* omission" (*ibid.*, 192-3) – and it is

[7] But see Phil Dowe (2000) for an interesting solution.

supposing the event away *qua* omission which is the only way to suppose the precautions were taken.

Lewis has left the issue of omissions as "unfinished business" – so too, shall I for now, while noting that supposing there is causation by omission introduces pressing questions which need to be answered. I will, because I am fairly confident that omissions can cause and be caused, assume that this matter is solved, however. (That is, at least until later, where I will spell out Schaffer's solution – which I accept for the duration of this paper.)

Accordingly, there is one specific case which contains omissions that I will make use of throughout my paper. This case is called "double prevention" (Figure 1). Double prevention, noticed first by Jonathan Schaffer (2000a),[8] has a structure that

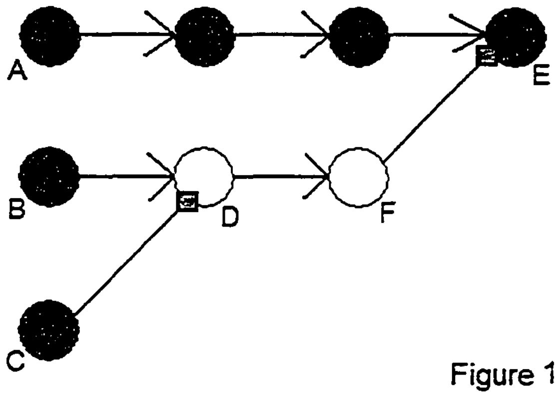

Figure 1

contains two causal preventions – hence the name. (These preventions are represented in the diagram by a straight line ending in a gray box. An event is represented by a single circle each of which, in the custom of Lewis, is usually called a neuron. That an

[8] Schaffer called these types of situations "disconnections."

event can potentially cause another event is indicated by a straight arrow leading from one neuron to another. When a neuron is filled, it is said to have 'fired' – in event language, the event is said to have occurred – and it will, in deterministic cases, cause the connected neuron(s) to fire – in event language, it will cause the event(s) to occur. Prevention takes precedence however, such that if a neuron, like D above, is prevented by one neuron, and stimulated by another neuron, that neuron will *not* fire. In contrast, consider the preventative line from F to E. Despite the presence of this preventative line, the E neuron still fires. This is so because F itself has not fired; had it fired, E would not have fired. For probabilistic cases, an event's probability is indicated by the value which is adjacent to its causal line. So, if the probability that C would cause D to fire was 0.5, this number would be written adjacent to the line connecting C and D. When there is a chance that an event will cause another event, I will make it clear, example by example, whether it succeeds to or not. I shall use this procedure for all the diagrams which follow, adding clarifications if necessary, along the way. Moreover, I shall number the diagrams sequentially; but, as a pedagogical aid, I will repeat old diagrams, and maintain their original numbering.)

The diagram of double prevention (Figure 1) is very abstract, but it is the template structure for real world examples. Consider: A plane (A) is on a direct path towards a mountain. If it does not swerve, it will certainly crash (E). An air-traffic controller recognizes the danger and tries to radio the pilot (B). Bonny, however, is a vicious vandal. She wants the plane to crash, so she places a bomb on the transmitting radio tower, destroying it (C). This prevents the transmission of the air traffic controller's warning (D), which in turn fails to prevent the crash by notifying the pilot (F). The pilot, unaware, flies the plane into the mountain (E).

A Lewisian analysis, using counterfactuals fits nicely with most people's intuition that Bonny is a cause of the crash. Had Bonny not blown up the tower, the air-traffic controller's message would have been transmitted, and the plane would *not* have crashed. (Applying Lewis method a little more stringently, the antecedent of the conditional $\sim C \square\!\!\rightarrow \sim E$ is true in the closest possible world; thus, the crash depends counterfactually on the bomb destroying the tower – and, thus, Bonny is a cause of the crash.[9]) The counterfactual analysis, like our intuitions, does not stumble over the troubling, but essential omission – occurring at neuron D – which results from the destruction of the tower.

I said above that the counterfactual analysis can give an account for the 'simple' causal situations (like situations when the president pushes the button and the world is destroyed) without trouble; this, however, is not entirely true. Equally as simple, but exponentially more difficult for the counterfactual theory to provide the correct answers, is a class of cases called 'preemption'. The cases within this class demonstrate that the counterfactual theory's focus on what goes on elsewhere in the world is too sensitive to the addition of non-causal factors to the original – 'simple' – causal line. Intuitively, any added events which are *extraneous* to the causal line should *not* affect

[9] An attentive reader will have noticed that I have tried to consistently say that C is a *a* cause of E, and not that C is *the* cause of E. In most cases, if not all of them, it would be erroneous to say that something is *the* cause of another thing. In this case, for instance, other causes of E might be the actual flight path of the plane and the persisting presence of the mountain. Note however, that the counterfactual analysis would not deny these claims. Surely it would assent to "if the mountain were not present, the crash would not have happened" and "if the plane had swerved, the crash would not have happened."

the causal status of that line – but, the counterfactual theory cannot speak in tune with our intuitions.

Let us consider a simple case first, and then show how small, seemingly irrelevant additions, change the verdict of the counterfactual theory. Consider Bonny the vandal again, on her own. She has decided to cause some destruction, so she throws a rock at a window, hits it and it breaks. When we ask, under the counterfactual theory, whether Bonny's throw is a cause of the breaking, we get the correct answer: had Bonny not thrown, surely the window would not have broken. So far so good.

Now however, consider the same case, but add Phil. The situation then, is as follows (Figure 2): Phil (A) and Bonny (C), both destructive vandals, are poised to

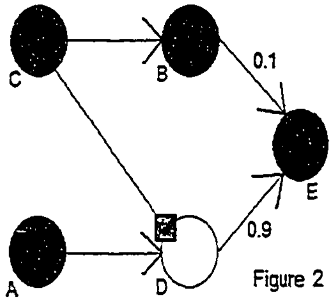

Figure 2

throw rocks at a window. Bonny throws her rock (C) (as in the original case), and Phil stands in waiting such that, had Bonny not thrown, Phil would have; thus, Bonny preempts Phil. (This situation is indicated by the preventative line with the grayed box leading from C to D, and is referred to as 'cutting.' Here, D is the the flight of Phil's rock through the air.) Add one further stipulation that Bonny is much less likely to hit the window – probability: 0.1 – than Phil, who has practiced quite a bit – probability:

0.9. Bonny's rock, nevertheless, does hit the window and break it. This situation is called "early preemption."

On Lewis' counterfactual account, Bonny does *not* come out as a cause of the window's breaking – despite the similarity with the case where Bonny was alone – because there is no counterfactual dependence between the breaking and the rock throwing; had Bonny not thrown, the chance of the window breaking would be higher because Phil would have thrown (and the probability that he would succeed is 0.9, which is higher than Bonny's 0.1). Thus, even though our intuitions tell us that Bonny's throw causes the window to break, the counterfactual analysis can not reach such a conclusion because her throw does not raise the probability of the occurrence of the effect.

Lewis suggests finding an intermediary event, B, such that the chance of the window breaking, E, counterfactually depends upon B, and that B in turn, counterfactually depends on the prior event C, Bonny's throw. The event B has to be late enough in the chain such that the alternate process (i.e, Phil's throw) has already been preempted.[10] On this modified analysis, we say that $C \Box\!\!\rightarrow B$, and $B \Box\!\!\rightarrow E$.

Lewis calls this type of counterfactual dependence a causal chain. Specifically, a causal chain exists for a finite number of actual events, C, D, and E if and only if D counterfactually depends on C, and E counterfactually depends on D. Using this

[10] I suppose that *strictly* speaking Figure 2 ought to have additional neurons between C and B, and A and D to represent the flight of Bonny's and Phil's rock (respectively). Presumably it is not Bonny's throw which prevents Phil, but rather Phil's sight of Bonny's rock in the air. There would still, however, be an intermediary event such that the effect depends counterfactually on it, and such that it depends counterfactually on C. Never mind; the case does not change – it is just clarified slightly.

definition, Lewis then redefines the notion of causation such that, C causes E if and only if C and E are actual distinct events, and there exists a causal chain between the events C and E.

This move takes causation to be the ancestral of counterfactual dependence. It is *necessary* to redefine causation as the ancestral of counterfactual dependence not

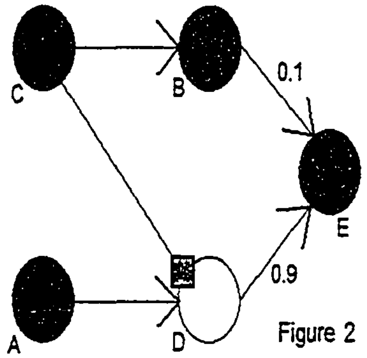

Figure 2

only to overcome the early preemption problem, but also to make causation transitive. For Lewis causation "must always be transitive" (1973, 200); and yet, as we can see in the early preemption case, when causation is defined purely in terms of counterfactual dependence, it is not. Consider Figure 2 again. While B counterfactually depends on C; and E counterfactually depends on B; it is not the case that E counterfactually depends on C, given that had C not fired, E still would have fired thanks to A: transitivity does not hold. Taking causation as the ancestral to counterfactual dependence (such that causation is a relation which holds between, e.g., A and C just in case there is counterfactual dependence between A and B, and B and C), is a way to ensure that causation will be transitive, *and* that the correct answer for early preemption cases can be provided.[11]

There is more than one way to cut a causal process, however (a fact which is often forgotten). And, if we change the way that the idle process is preempted, we change the causal structure significantly enough to undermine the 'ancestral' (chaining) amendment just given. The case represented in Figure 3 is structurally similar to Figure

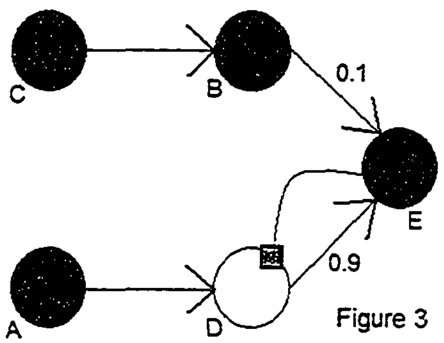

Figure 3

2, except that this time Phil (A) actually throws his rock at the window. Bonny (C) throws as well, but because she has thrown slightly harder (or slightly sooner) than Phil, her rock gets to the window first, breaks it, and thereby prevents Phil's rock from doing so. Let the previous chances remain the same: Bonny is much less likely to hit the window than Phil, but she, nevertheless, succeeds in doing so. This changed example is one of "late preemption," because the cutting occurs from the completion of the final event in the process started by Bonny (i.e, the firing of E; this type of

[11] It is not always the case, given that causation is the ancestral of counterfactual dependence, that where we have causation we will have dependence. Clearly, where we have dependence, we will have causation, but the converse does not hold.

prevention is indicated by a bent reverse preventative line from E to D). It is because the final event has actually occurred that Phil's action (the throwing of his rock) is causally idle.

Here, however, it is not possible to find a causal intermediary in the process C→B→E after the event which cuts the preempted line, because the event which does the cutting is the final event in the sequence. Thus, Lewis is forced to abandon his 'continuous' account of causation.[12]

This account gives way to another – quasi-dependence (Lewis 1986b)[13] – which is an attempt to exploit the obvious fact that (in late and early preemption) the cause is connected to the effect. (As Lewis later says, quasi-dependence was an attempt to appeal to "the intuitive idea that causation is an intrinsic relation between events" [2000, 184].) Even though quasi-dependence has detrimental shortcomings, such adaptations are testament to the strengths of considering the processes involved. That is to say, because all of the troubles in the problem cases arise from what goes on elsewhere in the world, an attempt which has 'narrower vision' will, perhaps, be more successful.

---

[12] Another problem that forces the abandonment of this analysis, which I shall not discuss here, is "trumping", first presented by Schaffer (2000b). Here, both lines run to completion – there is no cutting involved at all – but only one line has causal status. Schaffer's example is one of Merlin and Morgana; Merlin casts a spell at noon turning the prince to a frog at the following midnight. Morgana does the same, but at three o'clock in the afternoon of the same day. Because of the rules of that particular world, only the first spell cast each day will actually take effect. At midnight, the prince is turned to a frog by Merlin's spell, but because both spells actually reach the prince ,there is no cutting. (There is a question whether trumping is actually possible in this world; non-fantastic examples which do not rely on social norms are not forthcoming.)

[13] See section E - "Redundant Causation"

2.2. 'Local' Process-Theories of Causation

Process theories are properly deemed 'local' because they *deny* that what is important about any given causal situation is what goes on elsewhere in the world. Indeed, part of the motivation for local theories was to avoid such concerns; the local theorists, to be specific, wanted to focus on the connecting line (or lines) between events, in an attempt to provide the intuitively correct responses to the troubling preemption cases. Noticing that what caused all the troubles for the global theories was the presence of counterfactuals, the theorists were motivated, also, by the desire to develop a theory which could give an account of causation, while *avoiding* counterfactuals.

There are several flavors of process theories of causation that were developed in this light, all of which have the commonality that the cause is connected to the effect by some form of a line; they never make mention of what goes on elsewhere in the world. The connection that these types of theories focus on is usually cashed out in terms of a process. The theorists claim, that is, that the process between events is what determines the causal status of any particular event. Not just any process will do, however. There are plenty of processes in the world which are not themselves causal: A shadow moving along the ground (as the producing car drives forwards) is a process, but unlike the car, it has no causal status.[14] Any process akin to the shadow is called a

[14] Perhaps it is proper to say that the the shadow has no 'direct' causal powers. It could be that seeing the shadow causes someone to react in a certain way, but this – strictly speaking – is not caused by

pseudo-process.

One of the main differences between the diverging kinds of process theories is the means by which they distinguish a causal process from a non-causal – pseudo – process. Wesley Salmon (1984) gives an account which can be called a *mark transference* theory, where causal processes are only those where the (potential) cause is alterable at a local, single point in the process, and this alteration is maintained throughout the rest of the process. Thus, we can consider the flight path of the ball that breaks a window as a causal process, because at any point along its trajectory, we can mark the ball (e.g., by spraying paint into its path) at a single, local point, and this mark remains for the rest of the process.

The idea of mark transference is meant to supplement the basis of his theory, which is the idea of *causal interaction*. The definition of a causal interaction (given at Salmon 1984, 170), gives us the means to distinguish when a causal interaction occurs between two processes, as opposed to when it does not, by distinguishing the quantities that both processes would have maintained without an interaction from the quantities they would have after an interaction.

Salmon's account, however, has drawn fire. One of the aims of a local account of causation, remember, is to avoid the use of counterfactuals; it is not clear that Salmon's acount succeeds in doing so (see, for example, Hitchcock 1995). Clearly Salmon's definition of causal interactions does not avoid relying on counterfactuals: by

the shadow, but by the sight of the shadow. Compare this to a simple case of a car crashing into a wall: we can say that the car's crash caused the broken brick wall; but we do not wish to say that the collision between the shadow of the car caused the 'broken' shadow of the brick wall.

simply re-reading the above paragraph we can see that Salmon's definition relies on distinguishing the quantities that both processes *would have had* if they had not interacted from the quantities that they *would have had* if they had interacted. But even worse, as Hitchcock notes, Salmon's approach to differentiating pseudo from causal processes is counterfactual in nature as well. The idea of marking a process is a counterfactual question along the lines of "if I did such-and-such to this process, would it maintain the such-and-such?" Marking is not something that must be done; instead it is only a matter of whether it could be done. Thus, as successful as Salmon's theory is, it does not maintain one of its original motivations to avoid the use of counterfactuals.

Phil Dowe's (2000) conserved quantity theory, which was developed separately from Salmon's mark transference theory, is another attempt to give a local account of causation while avoiding counterfactuals. Dowe separates a pseudo-process from causal process by claiming that the latter "is a world line of an object that possesses a conserved quantity" (90). Other than saying that a conserved quantity is "any quantity that is governed by a conservation law" (*ibid.*), nothing more is given except that the best candidates for conserved quantities are those picked out by our current scientific theory (e.g., energy, momentum).

Dowe, using this method for determining which processes are the causal processes, then provides his theory of a causal interaction. A causal interaction, he says, is "an intersection of world lines [of objects] that involves exchange of a conserved quantity" (or more than one conserved quantity) among objects; where a world line is just the trajectory of an object in a space-time diagram; and where an object is anything that is scientifically defined or present in common day life (*ibid.*, 90). (Thus, he includes photons, waves, chairs, and people as objects.)

The upshot of this formulation, as the reader will notice, is that neither the definition of a causal interaction, nor the differentiating factor between pseudo and causal process, is stated in terms of counterfactuals. (These points make Dowe's account, in my opinion, the most successful local account of which I am aware.)

Despite their differences, both these accounts (and all the other local accounts) have in common the stipulation that actual flow of one "thing" across events is the deciding factor about causality. As such, local theories are tailor-made to handle the preemption cases which are a problem for the global theories such as Lewis'. Consider a brief application of Dowe's account, for example, to the situation which corresponds to Figure 3: Bonny (C) and Phil (A) both throw rocks at the window. Bonny throws her rock a tiny bit sooner than Phil, so her rock gets to the window first, breaking it (E). Because the window is already broken, Phil's rock passes through where the window once was; had Bonny not thrown, Phil's rock certainly would have broken the window. We can apply (the sketch of) Dowe's theory to see that Bonny's throw imparts certain quantities to the rock which it maintains through its flight; and, the rock's collision with the glass results in the change of certain related quantities in both the rock and the window. Because it never hits the window, Phil's rock is not able to transfer any conserved quantity to the window, and thus cannot be a cause of its breaking. (The same explanation will apply to the early preemption case of Figure 2.)

Thus, we can see that by exploiting the notion that a cause is connected to its effect (by some sufficient account of a process), early preemption (Figure 2) and the more troubling late preemption (Figure 3) are easily handled. Both are caused by Bonny, regardless of Phil the back-up, because of the connecting process between Bonny and the breaking of the window.[15]

No matter how we conceive the connecting process, however, our local theory will be plagued by the class of problems, presented above, called double preventions (Figure 1). The prevention – the non-occurrence of the event D – blocks any road we might travel down in trying to say that something – anything – flows across the events. The prevented event cannot transfer anything along the chain of events because there just is no event or object present to transfer a conserved quantity.[16]

Some theories, including Phil Dowe's (2000) and (2001) and David Fair's (1979), do try to allow for preventative causation like the kind found in Figure 1 (though only Dowe was actually aware of the double prevention cases, so far as I can tell), but doing so has – in both cases – called for the reintroduction of counterfactuals in some form or another. Thus, despite the strengths of theories which have a local character – at least one that sticks to its original motivations and remains 'pure' – they will be insufficient to account for *all* causal situations.

[15] Given a sufficient account of a process, the theory might even be able to avoid the trumping cases as presented by Schaffer. See note 12.

[16] A more in depth argument against local theories which takes this route is offered by Schaffer (2000a).

2.3. Prospects for a Theory of Causation

The considerations from §2.2 and §2.3 seem to point to the development of a hybrid theory, a combination of the local and the global theories. Given that both approaches on their own fail to accommodate *all* the troublesome causal situations at once, we ought to take the successful aspects of the global approach and combine them with the positive aspects of the local approach. Presumably, such a move would preserve the important capacities of each theory. With its global (counterfactual) aspects, the hybrid would be able to deal with the cases that contain omissions, such as double prevention; with its local aspect of a process, the theory would be able to bypass difficulties of (probability lowering) early and late preemption. Call such an approach, a *unified theory* – or *hybrid account* – of causation.

Of course, we need not take this route. Perhaps what is to be learned from all that has gone before is that there really cannot be a unified theory of causation because our world does not contain, at a *fundamental level,* one kind of causation. Perhaps, that is, we ought to infer that there are separate and distinct types of causation, each of which has its own separate theory. Call this the *dis-unified approach* to a theory of causation.

The dis-unified approach would come in at least two variations. The first, advocated by Chris Hitchcock (2003), would say that there are multiple, if not innumerable, distinct kinds of causation. Every causal situation that differs in some significant way will have its own theory of causation.[17] (Hitchcock does not spell out what differences would be 'significant' enough for any situation to count as a different kind of causal situation. For my own tastes, I would lean heavily on the respective structures of every situation, as presented in the neuron diagrams. The visual aids would help to demarcate differing structures and thus different situations requiring another account of causation.)

A second approach is to claim that there are two kinds of causation. This, itself, has two variations. The first variant (instantiated, for example, by Phil Dowe [2000]) holds that there is only one 'true' kind of causation, and the other is only 'ersatz'-causation. Typically, because we are aware from physics of the transference of quantities across events, the move is to say that the local theories capture the 'true' type of causation, while cases which contain omissions are merely ersatz-causation. Ersatz causation is then defined in terms of counterfactual instances of 'true' or actual causation.[18] The second variant (of which Ned Hall is a proponent) holds that there are two kinds of causal situations, *both of which* are genuinely causal in nature. Neither is more 'real' than the other, and each requires its own definition and explanation.

Though claiming such disunity is generally not favored – and requires solid reasoning for doing so – Hall's variant is simpler to defend than the Dowe's. Hall has to

---

[17] Hitchcock might be saying that there is an indeterminate number of causal situations. There might be three, or four, or five, or... (etc.), and we will only know once we have exhausted all the possible causal situations. It is in this light that I say there are only two different possibilities for the disunity of causation; I am subsuming the idea that there might be three kinds, or four kinds, or five kinds, ... etc., within Hitchcock's conception. (I think that what I say in §3.2 regarding the dis-unified theory's self defeating aspects will apply to all of these options *except* the 'infinitely many' approach.)

[18] See also David Fair (1979); Fair, however, does *not* state that the counterfactual instances of process causation are ersatz-causation.

provide justification as to why we ought to postulate a dis-unity at all – which is no easy task. However, an advocate of ersatz-causation like Dowe not only has to provide the same justification, but he must also provide justification for claiming that one type of causation is 'veridical' while the other is merely ersatz.

Given the additional hurdles for the second variant of the second approach, I shall not delve further (but see Dowe 2000). I will also quietly sweep Hitchcock's proposal of a vast disunity aside; for one reason, because it is relatively underdeveloped. Little has been said, for example, about specific definitions of causation, or when exactly a new theory of causation would be required. The second reason I sweep it aside is because, in my opinion, such a drastic disunity ought to be avoided if possible.

We appear to have two options then: either there is a unified theory to be had by combining specific aspects of the global and local theories; or, there are two differing types of causal situations, instantiated (perhaps) by the counterfactual theory on the one hand, and a process theory on the other. Given my aforementioned preference for unity, what reason do we have for advocating a dis-unified account? This is the topic of the next section, where we look at Ned Hall's work.

3. The Dis-unified Theory

3.1. Hall's Account

Ned Hall's (2004) analysis of causation begins by stating the "five theses" of causation which, intuitively, comprise the causal relation. He will, eventually, conclude that two of these theses conflict with the other three, and that accordingly there are two types (or kinds) of causation, which he calls *production* and *dependence*. The five theses are,

> 1) *locality:* causes are connected to their effects via spatio-temporally continuous sequences of causal intermediaries;
>
> 2) *transitivity:* if event C causes D, and D causes E, then C is a cause of E;
>
> 3) *intrinsicness:* the causal structure of a process is determined by its intrinsic non-causal character (together with the laws);
>
> 4) *dependence:* counterfactual dependence between wholly distinct events is sufficient for causation; and,
>
> 5) *omissions:* Omissions – failures of an event to occur – can both cause and be caused. (Hall 2004b, 225-6)

The theses (1), (2), and (3) comprise the *production* half of causation, while (4) and (5) – conflicting with the first three – comprise *dependence*. A paradigm case of production is one where, as the theses say, the cause is connected to the effect – e.g., a cue stick moving the cue ball, and the cue ball pocketing another; or, my throwing a rock, the rock hitting your window, and the window breaking. These are paradigm examples because they demonstrate the presence of all three theses: in the pool ball case, all of the events are connected (thesis 1); we say that the cue stick caused the pocketing of the second ball (thesis 2); and, we have no trouble saying (for example) that the reason the cue-stick caused the cue-ball to move is because of some intrinsic properties (or an intrinsic property) of the process.[19]

Strictly speaking, causal situations *qua* dependence, are those situations where the dependence and omission theses hold, (and where the first three *may* hold, but often times do not). A typical case of causation by dependence is one where an event of omission occurs. If you fail to pick me up (the event of omission), then I will be late for work; my being late for work depends on your not picking me up, because if you had picked me up, I would have been to work on time (barring some unfortunate accident).

But what grounds does Hall have for dividing up causation as he does? To begin, let's look at (1) locality. We can see straight off that it directly conflicts with thesis (5) omissions when we consider a case which is undeniably causal. Bonny shoots Phil. His heart is pierced by the bullet and he dies.[20] Phil's death, however, is caused by an omission because the piercing will *prevent* oxygen replenishment, which will cause oxygen starvation (i.e, the omission of oxygen in his blood). This, in turn, causes his brain-death by oxygen starvation. Yet, there are no spatio-temporal events between the heart piercing and brain death which can connect oxygen starvation to his death because the heart piercing prevents an event (the replenishment of oxygen) which

[19] Intrinsicness will be spelled out more clearly in what is to come.
[20] This is Schaffer's example (2000a).

would have prevented the death if it, itself, was not prevented. The non-event 'the failure to replenish oxygen' cannot possibly be connected to the event 'brain death' because the first is just that: a non-event. Such troubles lead to the conclusion that (1) locality conflicts with (5) omissions.

Hall believes that it will not suffice to say that the piercing of the heart is connected to the brain death by the series of absences between the shooting and the death. To say that a prevented event would have occurred in exactly the spot where it was prevented is quite a supposition: if a father grabs his child before the car hits her, it is a stretch to say that the child would have been hit in a specific spot X (very near where the child was grabbed) because it is to assume that the car – in the counterfactual situation – would not have swerved slightly, or would still have been going the same speed as in actuality, (or... etc.) all of which would affect the location of the accident.

Jonathan Schaffer, in his (2000a), gives just this solution to the shooting example presented above: the series of absences that culminate in the absence of oxygen can be located in the victim's lungs, his veins, his heart's ventricles, and in his brain (to name a few). Perhaps this response will hold for this particular heart piercing situation, but it is unclear if Schaffer is able to make the same move with respect to all prevented events: as Hall says, "Right now I am at home and hence I fail to be in my office; is the omission located there or here?" (2004b, 248). (The argument, as Hall states, can go many directions from here. However, I am of the mind – as we shall see in section 4 – that, perhaps, it might be alright to locate an omission in a bit of unfilled space-time, as

Schaffer insists.)

Locality also runs into trouble with (4) dependence. Strictly speaking, "*dependence* does not *quite* contradict *locality*, [but] it renders it satisfiable only by the most trivial laws (e.g., laws that say that nothing ever changes)" (Hall 2004b, 243). To

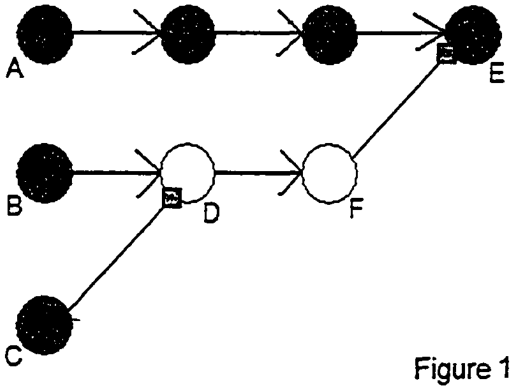

Figure 1

see this, recall the case of double prevention (Figure 1) – this time under a slightly different description. Bonny and Phil are on a bombing mission. Enemy (B) is on an intercept course with Bonny. Phil sees Enemy and fires on him (C), which prevents Enemy from continuing his flight path (D). During all this, Bonny continues her mission (A), and destroys the target (E). But, if Enemy had continued his flight path, he would have reached Bonny, and fired, (F) preventing her from succeeding. Now, our event E, depends counterfactually on C. *Ex hypothesi,* had Phil not fired, Bonny would have failed her mission thanks to Enemy. Thus, C is a cause of E. However, C can be classed as action at a distance, because, as above, we cannot – without difficulties – say that Phil's shooting Enemy is connected by absences to Bonny's bombing. Indeed, given that Phil shoots Enemy down some time before the bombing takes place, Phil's actions will count as actions at a *temporal* distance. Action at a temporal distance (and hence, action at a distance) amounts to C's non-locality with E because of the absences between them. And because the assumption that the locality *and* dependence theses are true drives us to accept that action at a (temporal) distance is possible (and even abundant) in nature, we ought say that if dependence holds then locality is trivialized.

Let us move on to the second thesis, transitivity, and its conflict with the fourth, dependence. One might raise an eyebrow in confusion if one were confronted with the seemingly ridiculous question "when is causation transitive?". However, serious counterexamples to the claim that causation is *always* transitive have been presented in some of the recent literature. What is common among all of these examples is not only the presence of at least one event of omission, but also a relation of dependence between events. Consider the following example, adapted from (Hall 2004a, 183):

> *Dangerous Business:* A man places a bomb at my door (C). That same evening my friend, a bomb expert, finds the bomb by chance and disarms it (B). The next morning I get up (survive) and eat breakfast (E).

Placing the bomb (C) causes my friend to defuse it (B). The diffusion of the bomb (B) causes me to survive and eat breakfast (E). So, if transitivity holds, then placing the bomb (C) is a cause for my survival to eat breakfast (E), which is clearly counterintuitive. Notice, too, that it is because of the thesis of dependence that we might be inclined to say this, given that B depends on C, and E on B.

To make the conflict between transitivity and dependence more apparent, and to highlight the presence of omissions at the same time consider yet another diagram (reproduced from Hall 2004b, 247): Figure 4. That C is a cause of E in this figure, is an absurd claim, because as Hall notes, "the little four-neuron network that culminates in F, is from the stand point of E, totally inert" (247); C never has a chance to cause E, because F can never fire. Yet, E counterfactually depends on F, which depends in turn on C. Both theses [(2) and (4)] cannot hold without absurdities arising.

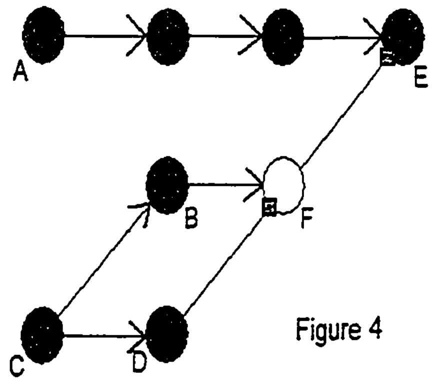

Figure 4

Notice, too, the presence of the omission F. If F had fired, then by all accounts, E would not have occurred. So, from the stand point of E, F's failure to fire – its omission – counts as one of the causes for E. And, we can notice that the omission F was itself caused by C. Again, by transitivity we get that C is a cause of E – but all that C has done "*is to provoke a threat to* [E]; [but] the very action that provokes the threat also manages to counteract it" (Hall 2004b, 252). If the omission thesis holds in all cases where the transitivity thesis holds, then there will be other conflicts of this type.[21]

Finally, consider thesis three, intrinsicness. Hall (2004b) formulates it as follows:

> "Suppose an event E occurs at some time t'. Consider the structure of events S that consists of E, together with all of its causes back to some arbitrary earlier time t. Then, any possible structure of events that exists in a world with the same laws, and has the same intrinsic character as S, *also* has the same causal character, at least with respect to the causal generation of E." (244)

What has been offered here, is a general rule which we can apply when we assume that the intrinsicness thesis is true. That is, when we assume that what makes a causal situation are those factors which are intrinsic to it, we can apply the rule – given above – to determine for any given situation, if an event is the cause of another.

Briefly then, let us assume the intrinsicness thesis to be true and see how it can conflict with the dependence thesis. By the definition of intrinsicness we can judge a structure of events S' to be causal by finding a match with another similar structure of events, S (which is causal by assumption). Now, reconsider that double prevention case in Figure 1. We might be inclined, in keeping with the intuition that Phil (C) is a cause for the bombing (E), to find an event structure S which includes E, back to the

[21] Note, however, that the theses of omissions and of dependence do not *always* have to conflict with transitivity. If you fail to pick me up for work, that will cause me to be late; and, my being late for work can cause me not to get to drink coffee in the morning; and, it is certainly not absurd to say, with transitivity, that your failure to pick me up caused me not to get to drink coffee in the morning. The point, instead, is just that there can be conflicts among these theses, and that the truth of the last two (dependence and omissions) does not necessitate the truth of the second (transitivity).

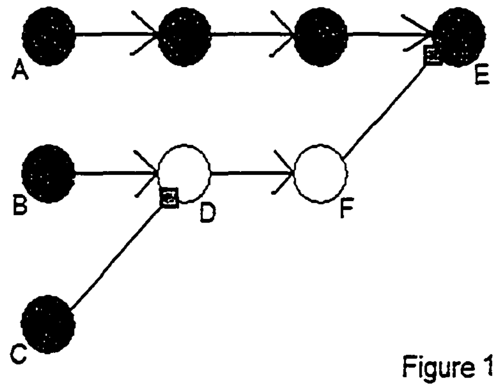

Figure 1

time of C. Starting from the bombing, the relevant (salient) events appear to be all those that constitute Bonny's flight path and button pushing. By dependence, as we have already seen, we must also include Phil's actions, which probably includes his spotting Enemy and firing on him; the bullets hitting Enemy's plane; and Enemy's plane blowing up. Let's say that all of these Phil-events and the previous Bonny-events comprise our structure S.

Now, consider the same case, but add the stipulation that Enemy would *not* have fired on Bonny unless given orders from his home base. Further, Enemy's home base never had the intent of ordering the attack on Bonny: Enemy poses no threat to Bonny at all. By firing on Enemy, however, Phil irritates Enemy's home base, causing them to order all of their fighter pilots in the area to shoot down Bonny. As it happens, Enemy is the only one in the area, so no one else poses any threat to Bonny. In this slightly changed case, all of the events from S are present: we have Bonny's flight path, her dropping the bomb, Phil spotting Enemy and shooting him down, etc. That is, there is a structure of events S' within the new case such that it matches our structure S from the original situation exactly. But for S', it would be absurd to say that Phil is a cause of the bombing: It is only because Phil fires on Enemy, that Enemy poses any threat at all to Bonny. Phil is *not* a cause of the bombing in the new case, and our definition of intrinsicness is violated. So, we must conclude that our original structure S was incomplete, and modify it to include other events – such as Enemy's orders. Making this modification, which is necessary to explain the causal situation and maintain the intrinsicness thesis, is like admitting that these new events are causes of the bombing, which is "ridiculous" (Hall 2004b, 245). (In both cases, Phil's actions are what Hall calls "extrinsic presence of enabling factors[... ;] enabling in the sense that if they were absent, there would be no dependence of the bombing on [Phil's] actions." (245).)

"Notice ... that it is exactly the inclusion of [Phil's] action as part of the causal history S that is the culprit: Once we include it, we must also include (on pain of denying *Intrinsicness*) all those events whose occurrence is required to secure the counterfactual dependence of the bombing on this action" (Hall 2004b, 245). When we assume that both dependence and intrinsicness are true we are driven to greater and greater lengths to make them coincide. If we want to maintain the truth of dependence, then we have to include Phil as a cause of the bombing. Yet, if we want to keep intrinsicness true, then we need to add even more factors – like Enemy's orders from his home base – which must also count as a cause of the bombing. Thus, these two, dependency and intrinsicness, conflict.[22]

[22] Hall also introduces the idea of "extrinsic absence of disabling factors" which are "disabling in the sense that if they were present, there would be no dependence of the bombing on [Phil's] action"

Let me present yet another neuron diagram, Figure 5, to illustrate the conflicts between the omission thesis and the intrinsicness thesis.[23] (Notice in this diagram that D, with its slightly thicker border, is a stubborn neuron: it needs to be stimulated *twice* in order for it to fire.) If we, again, assume the intrinsicness thesis is true, then we might try to capture the causal structure of Figure 5, by including all the positive events in our structure S. And, in keeping with the omissions thesis, we might be inclined to include the omission of B as one of E's causes, given that had it fired (and

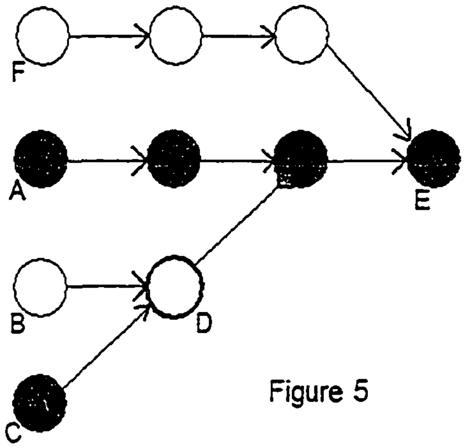

Figure 5

C still fires), E would not have fired. So, let us say that the causal structure of S consists of B's failure to fire, and also all of the positive events (C and A→E), back to an earlier time. Now, assuming intrinsicness to be true means that for any situation where we can find a structure S' that matches the structure S (taken from Figure 5), we

(245) - such as a situation where there is a bomb under Enemy's seat which would have gone off if Phil had not fired. I am not entirely convinced that this situation is an example of the conflict between dependence and intrinsicness, however, because it requires the admission that Phil, in the changed example is *not* a cause of the bombing: an intuition which I do not have.

23 This diagram is Hall's (2004b, 250).

will know that S' is causal as well; and, thus we will know that B's failure to fire is a cause of E.

But, trouble starts right away because it is not the case that every time we find an S' which duplicates S, that B's failure to fire will count as a cause of E. Consider, for instance, a causal structure exactly like the one depicted in Figure 5, *except* that F fires at the same time that B fails to fire (and A and C still fire). In this new case, the neurons running from F to E will all fire and the line will run to completion. In this case, it is not clear at all why we ought to count B's failure to fire as a cause of E, for even if B had fired E would still have occurred. This verdict (that B's failure to fire is no longer a cause of the firing of E) may be prevented simply because of the line that begins at F, but we can find in this new example, an exact duplicate of S can be found – so intrinsicness conflicts with omissions. (F, in this second situation, is yet another case of the extrinsic presence of enabling factors.)

These considerations conclude the arguments presented by Hall. They motivate him to separate the causal relation into the two types (as I have mentioned): *production* and *dependence*.

In terms of actual definitions for each type of causation, Hall believes that half of his work has been finished for him by Lewis; until more difficult circumstances dictate otherwise, Hall believes that we can just understand *dependence* in the normal counterfactual way spelled out by Lewis (e.g., the counterfactuals are of the ancestral, non-backtracking variety, where it is only the truth of the counterfactuals at the closest possible worlds which matters, etc.).

Production, Hall believes, is best captured by some version of the intrinsicness thesis. It will not serve any great purpose if I am too specific about his program here, but the briefest possible of outlines ought to be given. I have already intimated that the definition of intrinsicness involves some form of duplication. For Hall, this duplication is of the structurally non-complex, primitive cases. For instance, preemption cases – of the early or late variety – contain a structurally simpler version where Bonny is alone and she throws her rock at the window. From this, we can extract what Hall calls a blueprint. Call this blueprint, S. When we add elements to this blueprint – like Phil's throw – what we are doing is adding things which are extrinsic to the original causal process.

These additions have no (relevant) effect on the status of the Bonny-line in so far as it is considered causal. That is, for cases of preemption we can determine Bonny to be the cause of the breaking because of the presence of a structure, S', which matches S in the relevant way. (S of course, was obtained from the structurally simpler case where Bonny was alone.) The structures match *in the relevant ways*, and not exactly of course, because when we take the primitive case where Bonny is alone and then add Phil's throw, we ring very minute changes throughout the example (say, e.g., from the small gravitational effect that Phil's rock has on Bonny's) which prevent the cases from having exact similarity (Hall 2002b, 287).

There are very many, if not infinite, numbers of ways that we can insignificantly change the causal structure S – adding Phil's throw is just one way – without affecting the causal status of Bonny's throw. For all of these cases, our result ought to be the same: Bonny's throw should come out as the cause of the window breaking. In this light, Hall proposes that we treat this collection of (insignificantly) changed situations as the "blueprint-class" (*ibid.*, 289). Every variation of S in this blueprint-class will contain all and only the events between two times, t_1 and t_2 (the time of the cause and the time of the effect) which are uniquely sufficient for bringing about the effect in question (e.g., the window breaking). Every causal situation which is an instance of production will contain some structure S which belongs to a blueprint-class (or which serves as the structurally primitive situation from which we extract our original blueprint), and this class determines one way or the other, the causal status of the event in question.

Before moving on to an analysis of Hall's theory, let me briefly reiterate his views. I have recounted Ned Hall's arguments for the disunity of causation. He claims that causation comes in two kinds, *dependence* and *production.* Causation, *qua* production, is captured by three (of the five) causal theses: transitivity, locality, and intrinsicness. Transitivity is straightforward, where if event C is a cause of D, and D is a cause of E, then C is a cause of E; the locality thesis holds that causes are connected to their effects via spatio-temporally continuous sequences of causal intermediaries; and the intrinsicness thesis states that the causal structure of a process is determined by its intrinsic non-causal character, together with the laws. Assuming the truth of the intrinsicness thesis usually entails a procedure of duplication of the process in question to determine its causal nature. The other half of causation – causation *qua* dependence – is captured by the other two of the five theses: dependence and omissions. The first says that counterfactual dependence between wholly distinct events is *sufficient* for causation; the second says that omissions – failures of an event to occur – can both cause and be caused. Using these, I have shown how the attempts that Hall makes to achieve a hybrid account of causation (where the local aspects and the global aspects were considered to be independently sufficient for causation) results in the conflicts between the respective theses. These results have motivated him to postulate the two aforementioned types of causation.

3.2. Analysis of Hall's Account

Dependence and production, according to Hall, jointly exhaust the causal situations for which a theory must provide the correct results. Hall's account of production handles the 'local' preemption cases nicely, while his account of dependence provide answers for the 'global' double prevention case. (These cases are represented by Figure 2 and 3, and Figure 1, respectively.) And yet, because all causal situations, so I would have you believe, can be represented by the neuron diagrams, a potential problem looms in the darkness. It seems at least intuitively plausible that we might be able to combine a local preemption problem with a global double prevention problem, to form a whole new causal situation. It is not (*prima facie* at least) clear how a theory such as Hall's would handle such a causal situation. Given that preemption cases are handled by production and double prevention cases are handled by dependence, what Hall would be able to do with the new problem that is a combination of problem types is an interesting question. Before we get to this, however, let us briefly turn to see how Hall's two kinds of causation deal with the problems of preemption and double prevention.

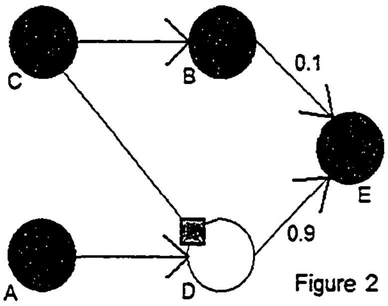

Figure 2

Production is ideal for giving the intuitively correct answer for all forms of preemption. Recall early-preemption (Figure 2) – this time under a new guise. Phil fixes to blow up a radio tower by placing a puny bomb on it (C). Bonny, also fixes to blow up the tower (A) with her huge bomb. Given the size of her bomb she very likely would have succeeded (0.9). Nevertheless, she fails to plant her bomb because she is overcome with laughter at the sight of Phil's puny bomb. Phil's puny bomb blows up the tower (E), despite its very low chance (0.1) of doing so.

Implementing causation *qua* production, we first designate the entire preemption structure as S'. Then, we ask if there is a structure within S' that belongs to a larger blueprint-class. To coincide with our intuitions, the structure we seek would have to be one where the Phil line is considered causal.

We are in luck, of course, because the structure which we seek is simply the one where Bonny is not present at all. Thus we have Phil placing his tiny bomb on the tower, and the tower subsequently being destroyed. Call this structure S. In S, Phil is undeniably the cause of the tower's destruction; and, fortunately, within the preemption structure S', we can find a near perfect match with S.[24] S will surely belong to some blue-print class where C (Phil's bomb) counts as a cause of the tower's destruction, so we are justified in the preemption case, in deeming C the cause of E. I leave it as an exercise for the reader to see that the same procedure can be repeated in order to solve the late preemption case.

The problems that Hall proposes are best dealt with by dependence are the cases of double prevention.[25] Again recall the fighter pilot example (Figure 1):

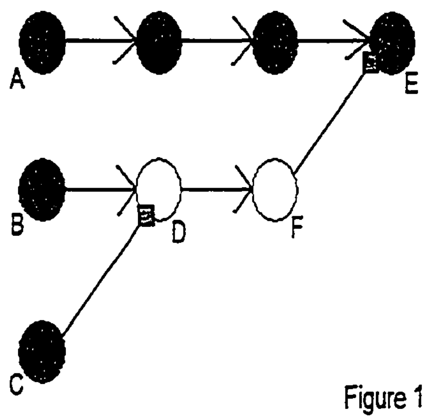

Figure 1

Phil and Bonny are fighter pilots on a bombing mission. Keen-eyed Phil (C) spots Enemy (B), veers left and shoots him down; this prevents Enemy from continuing his

[24] I say "near perfect" because presumably the presence of Bonny in the preemption case could make some minute difference to S in the light reflection, the direction of stray photons, etc.

[25] Actually, this is what Hall *should* have said, but does not. For these types of cases Hall erroneously implements production. However, by using intrinsicness Hall is forced to deny that C is a cause of E; I do agree with this conclusion. More on these issues presently.

flight path (D), and shooting down Bonny (F). Meanwhile, Bonny (A) continues on her way and destroys the target (E). Recalling that Hall simply adopts the Lewisian form of counterfactuals for his theory of dependence, we must ask whether E causally depends on C. Here, our intuitions (generally) conform: If Phil had not shot down Enemy, (by stipulation) Enemy would have shot down Bonny and she would have been unable to finish her mission; thus, the bombing counterfactually depends on Phil's action.

But now let us turn to the trouble which lurked nearby in the darkness: the combination of preemption and double prevention. Let us call the new situation, keeping in tune with Schaffer and common sense, a case of *preemptive double preventions* (Figure 6*).[26] We can see very vividly from the diagram, that the case is a

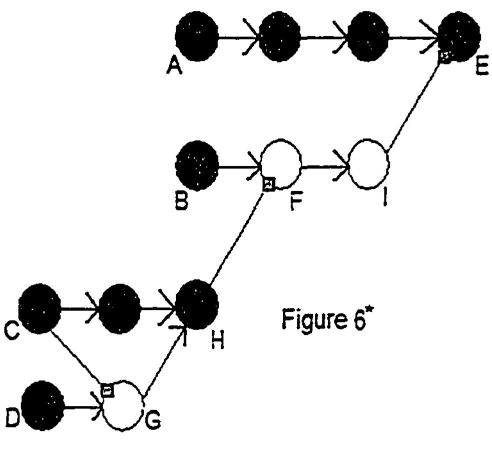

Figure 6*

mere concatenation of the local and global problems. The double preventions remain the lines between H and F, and I and E; while the preemption remains the link from C to G.

This situation, however, need not remain an abstract conglomeration of neuron

[26]Hall has left these cases as unfinished business.

diagrams – it can be put into (semi-)realistic terms. Fighter pilots Phil and Bonny are on a bombing mission one more time, but now Phil is accompanied by Wingman. Phil (C) spots Enemy (B), veers left and fires at enemy, shooting him down (H). Waiting in the wings is Wingman (D), who would have veered left (G) and shot down Enemy if Phil had not. Phil shooting down Enemy, prevented Enemy from continuing his flight path (F) and shooting down Bonny (I), which would have prevented the bombing. Nevertheless, Phil *does* shoot down Enemy, so Bonny (A) is able to complete her mission: the target is destroyed (E).

Intuitively (for most at least), Phil is still (as in the double prevention case of Figure 1) a cause of the bombing. Yet, for reasons given in §3.1 concerning the conflict between the intrinsicness thesis and absences (which occur, here, at F and I), Hall is unable to implement his theory of production to solve the case. Nor, moreover, is causation *qua* dependence – at least as Hall originally presents it – capable of determining that Phil is a cause of the bombing. D's (Wingman's) presence botches the verdict that E (the bombing) depends on C (Phil's action), because without C, D would have prevented F (Enemy's continued flight path); preventing F stops I from being able to fire (shooting Enemy down stops him from being able to attack Bonny); and, since I does not fire, it cannot prevent E (the bombing), so E still occurs. Thus, E would still have happened if C did not happen (thanks to D), *contra* the counterfactual requirement.

Hall notes that this shortcoming is a result of interpreting the counterfactual as Lewis proposed, and suggests reinterpreting them. He briefly mentions, but does not develop the idea that in Figure 6 we could hold other specific aspects fixed (not just everything but the cause) when we evaluate the counterfactuals. This point is strengthened by juxtaposing Figure 1 with Figure 6; in a similar way in which we consider Phil a cause of the bombing in the Figure 1 case, we consider him a cause in the Figure 6 case. I wish to try to fill out some of the details of this approach to see whether it would be lucrative in handling preemptive double preventions.

There is some preliminary business to handle first, however. I have taken certain liberties with Figure 6*, in order to make apparent the relation this type of causal situation has to the more 'primitive' cases as seen in Figure 1 (double preventions) and Figure 2 (early preemptive prevention); that is, I have 'fudged' a bit to make it apparent that this case is a mere concatenation of the two figures. A scrutinizing eye, however, will have noticed that it is not exactly accurate to have Phil's

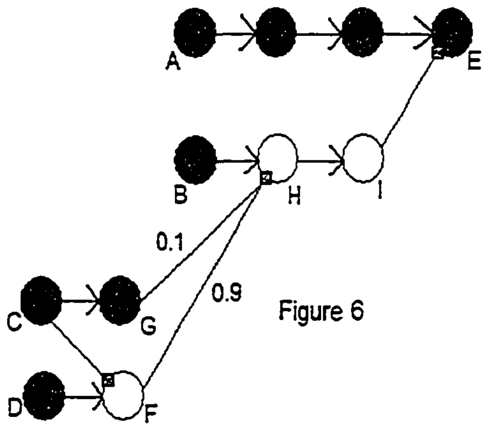

Figure 6

line and Wingman's line converge on H, if H is understood as a firing-on-enemy event. This is because Wingman's firing would not have been the same event as Phil's firing, and thus, should not have the same neuron in the diagram. The diagram should have looked like (this new) Figure 6. This change is insignificant to the case, however, except that the structure is *less obviously* a chaining of the two primitive cases. There is no major difference between Figure 6* and Figure 6: like the event F from Figure 6*, we have an event (H), such that if it were prevented, it would prevent a process (B→H→I→E) from preventing another process (A→F→E). G and F, from Figure 6, respectively, would be Phil and Wingman firing on Enemy, which would serve to prevent H. Either G or F could have prevented H from occurring (and, G in fact does prevent H).

That we can chain 'early-preemption' with 'double prevention' should hint that we can also chain the latter with 'late-preemption' as well; we would get Figure 7. The

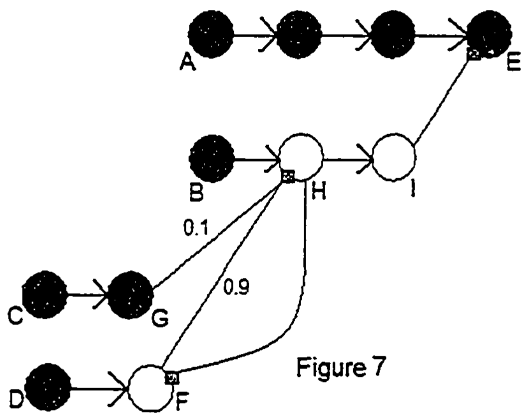

Figure 7

description of such a case would be the same as the description for Figure 6, except that this time, both Wingman and Phil fire, Phil slightly sooner than Wingman. Phil's bullets get to Enemy first and destroy him, which preempt (ever so slightly) Wingman's bullets from doing the same thing. Because Enemy is destroyed, Bonny is able to carry out the rest of her mission so the target is bombed.[27]

What then of Hall's suggestion to hold certain elements fixed, for preemptive double preventions?[28] Let us simply adopt Hall's suggestion: In Figure 6, Wingman does not fire on Enemy, so if we were to hold this fixed when considering the truth of the counterfactual, then Phil will turn out to be a cause of the bombing. As a general rule of thumb, we can state that in evaluating the counterfactual we ought to hold fixed (in some relevant way) the "pure dependence preventer" (Collins 2004, 111). Pure dependence preventers are those events whose presence prevents the effect from coming out as counterfactually dependent on the intuitively correct cause. This type of move leaves much to be done from case to case, because it is up to the evaluator to determine which event is the pure dependence preventer. Nevertheless, the work is not hard; often it is quite clear which event exactly is the trouble-maker. In Figure 6, for example, it is the presence of Wingman which prevents the bombing from depending on Phil's actions.

---

[27] Perhaps the most crucial objection to both of the combination diagrams, concerns the reverse-preventative link from H to F. Technically speaking because H does not fire, the prevention back to F can not take affect (and thus F should have fired). For this diagram then, I shall stipulate that if H is prevented its reverse-preventative link will take affect. This is not out of tune with the example: it is because Enemy is no longer flying that Wingman's bullets are preempted. It is only because the neuron diagrams are so rigidly drawn, while the descriptions of the events that accompany each neuron are significantly less rigid, that this objection even surfaces. Perhaps with different descriptions this objection could even be avoided. Nevertheless, the most that raising this objection can achieve is to highlight a possible shortcoming of the structural diagrams– which is, for all intents, minuscule.

[28] Hall takes issues with slightly different cases that cause trouble for the suggested modified account of counterfactuals. The case types, which are typical problem causers for this suggested analysis, are ones similar to when the preempted backup is, *prima facie*, less fallible than Wingman. One such case is where Wingman is replaced by an invulnerable shield around Bonny. Hall is unsure how to differentiate between the Wingman case and the Invulnerable shield case – but I do not think it is that hard. The invulnerable shield case is really structurally different than the Wingman case, because unlike the Wingman case, Enemy never really has a chance to shoot down Bonny. Thus, the all-important preventative link (from Figure 1) between F and E would be lost. The collection of Phil-Enemy neurons, in the invulnerable shield case, would then not be connected to the A→E line.

The trouble with advocating such an approach, however, is that the approach will likely oppose any attempts that Hall makes to defend his account of production, and more generally his reasons for positing two kinds of causation. Figure 6, difficult as it is, is merely the concatenation of the two cases given above in Figures 1 and 2. That is to say, it is the combination of the preemption problem – which was dealt with by invoking the production model – and the double prevention problem – which was dealt with by invoking the dependence model. But, if Hall succeeds in giving the intuitively correct answer by using differently interpreted counterfactuals, then it is not clear why he needs the production type of causation at all. What Hall will have done by applying the improved interpretation to Figure 6 is to have, at the same time, given an answer for the preemption problem. In fact, it is just the presence of the preemption in this new case which prevents Hall from using the original counterfactual analysis as posited by Lewis. If Hall's modified account of counterfactuals can solve the preemptive double prevention problem, it will have solved the (more primitive) preemptive problem as well. To see this, let us bear in mind the new account of counterfactuals as I have sketched it here, and apply it to mere preemption.

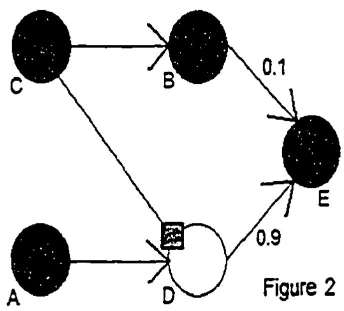

Figure 2

Consider the case from Figure 2 again: Bonny (A) and Phil (B) both plan to plant bombs on the tower. Bonny fails to plant her bomb because she is overcome with laughter. Phil plants his bomb and the tower is destroyed (E). Now, using Hall's new approach of modified dependence, we hold fixed that Bonny does not actually place her bomb on the tower and then evaluate the counterfactual. Here, Phil comes out as a cause because the right counterfactual relationship occurs; had Phil not placed his bomb on the tower, the tower would not have been destroyed (holding Bonny's non-act fixed).

As is usually the case, late preemption of Figure 3 proves to be too difficult for an amendment so simple as holding the pure dependence preventer fixed when we analyze the counterfactual. Because in late preemption the backup process actually runs on route to the final event, and thus the initiating event occurs, we cannot just hold the dependence preventer fixed. Bonny and Phil both throw rocks at a bottle, but Bonny throws sooner, so her rock gets there first and breaks the bottle. If we evaluate the counterfactual "if Bonny had not thrown, the bottle would not have broken" while holding the fact that Phil threw fixed, the counterfactual will, of course, come out as false. So, this account is not sufficient for solving late preemption.

Thus, it might seem that because of what has transpired in the previous paragraph, Hall's modified account will handle the new preemptive double prevention cases, without being sufficient for the more primitive preemption cases which would be handled by production. However, in actuality, the new account of dependence is not even sufficient to solve all variations of preemptive double preventions. For example, let us say that Wingman does fire on Enemy, but that Phil fires first (or that Wingman fires a slower missile, while Phil uses his machine guns). The modification changes the case to be a true late-preemptive double prevention case (as in Figure 7), and the new dependence account will fail to provide a solution, *for exactly the same reasons* that it failed to answer the more primitive late preemption case. Thus, because the new dependence account does not mesh with our intuitions for *all* variations of the preemptive double preventions, it must be abandoned.

Hall, however, can just go one step further in order to solve both variants of the the combination cases. He can stipulate not that we hold the pure dependence preventer *fixed*, but that we stipulate it *away*. This move would be similar to the definition provided by Collins (2004), who claims that,

> "C is a cause of E iff there is a chain of counterfactual dependence linking C to E, or there would be such a chain were it not for some pure dependence preventer." (111)[29]

Moreover, the move would simultaneously handle the early and late preemptive double preventions. Applying this method to both would produce the conclusion that Phil is a cause of the bombing because, when the counterfactual is evaluated we suppose that the back up process (Wingman's process) is not present. Under these conditions, regardless of whether Wingman actually fires on Enemy or not, Phil comes out as a cause.

Unfortunately, the successes of the new dependence account in the early and

[29] This account is abandoned by Collins because it will not give the correct answers for trumping preemption.

late double prevention cases also extends to the more primitive cases found in Figure 2 and 3. For example, when Bonny throws slightly sooner than Phil, breaking the bottle before him, we can determine her to be a cause of the bottle breaking because during our evaluations we suppose away Phil's throw (and perhaps even Phil) entirely. The application is similar for early preemption.

Providing the correct solution to preemption problems with his new dependence account is not at all desirable for Hall because it raises questions (that he *must* answer) about the necessity of his account of production. The cases of preemption separate themselves as causation *qua* production because they adhere to the intrinsicness, locality, and transitivity theses – which directly contradict the dependence and omissions theses. But, if we can handle the preemption cases with a modified version of dependence, then it is quite likely that we can handle *all* causal situations with the same modified version of dependence, including the cases where where Hall wants to use production. That is to say, the distinction of two types of causation becomes unnecessary with the development of this new kind of dependence, because it solves all the troublesome cases. Causation *qua* production is a superfluous supposition if Hall's new approach of dependence succeeds for cases of preemptive double prevention, because it will succeed for cases of preemption as well.

If I am mistaken in this criticism of Hall, however, and he is not subject to a self defeating position, he still has another very pressing question to answer. Pretending that this last issue is overcome, Hall has not given us a clear explanation of when we ought to use one type of causation instead of the other. In fact, if he is leading by example, it appears he may lead those who adopt his approach well astray. I said a while back that Hall analyzes the case of double prevention (Figure 1) in terms of the dependence half of causation; this was incorrect. In actuality, Hall uses production (2004b, 263).

Hall, correctly, considers A to be a cause of E. Applying Hall's account of intrinsicness, we can imagine a simpler situation which involves *only* Bonny and the target. From this situation we can identify a (rough) structure of events, S, starting at some time t which consists of Bonny's flight path, the persistence of the target, the pressing of the button to drop the bomb, and the destruction of the target (which occurs at a later time t'). Now with S in mind, we can look at Figure 1 again. We are able to find a duplicate structure, S', of the structure that we identified in the simpler case, thus we are able to say that A is a cause of E. Because we will not be able to find, in a simpler case, an event structure S" which would include the occurrence of C, we *cannot* count C as one of the causes of E. (The same can be said for the cases in Figure 6 and Figure 7. In both cases Phil cannot be considered a cause of the bombing because his actions are extrinsic to the process of Bonny's flight and the site's destruction.)

Denying that C is a cause, however, is highly counter intuitive. To begin, these cases were raised by Schaffer (2000a), to demonstrate the inadequacy of local theories (such as Hall's version based on intrinsicness). And, even if we were somehow to overcome our intuition that Phil's action (C) is a cause of the bombing (E) in this particular case, it would be even more difficult to do so for the other cases of double prevention which are ubiquitous in nature. Consider the heart piercing example I gave in the last section: Shooting someone through the heart is surely the cause of that person's death. However, this situation is a case of double prevention, very slightly modified (Figure 8).[30] The bullet pierces the heart (C). The man breaths in (B). The

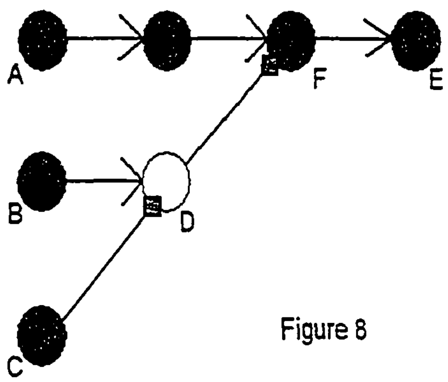

Figure 8

bullet prevents the resupply of oxygen (D). The resupply of oxygen would have prevented oxygen starvation (F), which in turn would have prevented the process A→E (where A is the depletion of oxygen in the brain; F is oxygen starvation in the brain; and E is brain death) which leads to the death-event (E). Hall may be inclined to say that Phil does not cause the bombing (in the original Figure 1), but then he has to do the same for shootings which cause death.[31]

Hall might be inclined to say (though he does not actually do so) that while A *produces* E, E still *depends* on C: thus, C is a cause of E. In my view, this move is

[30] The changes that I have made to the double prevention case have shortened the B-line. Nevertheless, there are still two preventions (leading from C to D and D to F), so the case is only insignificantly unchanged.

[31] For even more cases that most would consider regular every-day causal situations but what are actually double prevention cases, see Schaffer (2000a).

rather *ad hoc.* Because we can count Bonny's actions, as well as Phil's actions, as a cause of the bombing by implementing dependence alone – after all, if Bonny had not dropped the bomb, the target would not have been destroyed – it is unclear (at best) why Hall would use production *and* dependence, when just one of the theories is sufficient.

Moreover, when he switches to the more interesting and troubling case of preemptive double preventions, it is not at all clear what justification he has for using his theory of *dependence.* If he is set on using production for double prevention, what is it about the preemptive double prevention case that justifies switching to his global theory? Surely it is not the introduction of a problem (i.e, early/ late preemption) that a local theory is tailor-made to solve!

The murky water here is further muddied by Hall's mistaken attempt to handle a prevented switch with his theory of production (2004b, 264). A switch (Figure 9) is

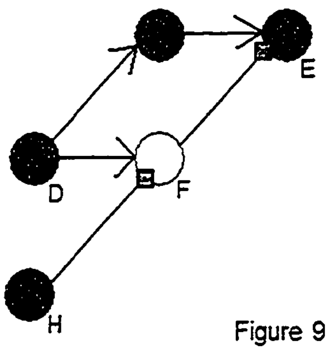

Figure 9

something which both causes a threat for some event, and causes a neutralizer for that threat at the same time.[32] Crucial to the occurrence of E here, is the firing of H: had H

[32] Compare Figure 4 for another instance of a switch.

not fired, F would have; and the firing of F would have prevented E. But, by using the intrinsicness thesis (i.e, causation *qua* production), Hall eliminates the possibility that H can count as a cause because of his own arguments for the contradiction of omissions and intrinsicness. And yet the situation is easily dealt with by the dependence half of causation: had H not occurred, E would not have occurred.[33]

All of this is indicative of a larger problem that any theory which posits a disunity must address. It is not clear, given that we have working accounts of the two types of causation, when one account rather than another is the definition of causation that we ought to employ; if we do not know the answer to this, then we cannot know which theory ought to be implemented. Notice, too, that this is not a trivial issue, given the drastically different results that each theory gives for the case of double prevention. Moreover, given that the structure of double preventions really is quite common, showing up in cases where we least expect, the issue is even more poignant.

To sum up, Hall's approach suffers from the problem which lurked in the dark – the combination of preemption and double preventions to form 'preemptive double preventions' – which seems to be an insurmountable, critical problem for any theory which would try to separate causation into more than one kind.[34] Moreover, it appears to suffer from an epistemic barrier which surfaces in questions concerning what type of causation a specific situation might be. When we are unsure which type of causation

[33] Hall is able to add restraints to his theory of intrinsicness, limiting again which 'simple' structures we are able to draw from (see 2004b, 264-5), but it not clear why we ought to bother with these changes, given that the case is easily dealt with by dependence.

[34] It may not be an insurmountable problem, though, for the theory that uses a new theory for *every different* causal situation. Such a discovery would just warrant another theory entirely. What that theory might be is up for grabs, of course.

we are dealing with, we cannot know which theory to use. And, finally, Hall still has the unanswered question – even when there is no epistemic barrier at all – of when one type of causation ought to be used over the other. He seems to lack a clearly (or accurately) drawn line.

4. The Unified Theory

4.1. Schaffer's Account

Even though Hall's account seems self defeating, his arguments for the incompatibility of the five theses which comprise the causal relation still stand. It seems we are left with nothing: any hope the dis-unified account might have had is dashed by combining the preemption problem with the double prevention problem (to get preemptive double prevention); and, at the same time, the dis-unified theorist presents compelling reasons for accepting that very dis-unity, and rejecting even the mere possibility of a unified theory.

Nevertheless, in the face of these criticisms, Jonathan Schaffer seems to have developed a unified theory which manages to handle (most of) the troublesome cases, preemption (both early and late) and double prevention. Schaffer recognizes the strengths of each type of theory – global and local – and that they are tailor-made to handle certain types of cases. Accordingly, he suggests that we ought just to combine the front running local theory, with the front running global theory.

This approach works, as we shall see, because the inconsistencies between the theses (as discovered by Hall) are inconsistencies which hold only under very specific formulations of those theses. Recall that Hall's unified approach was an attempt to combine the 'global' aspects (dependence and omissions) with the local aspects (locality, transitivity, and intrinsicness) under the assumption that each thesis *on its own* was sufficient for causation. Schaffer's unified theory, however, does *not* stipulate that the global and the local aspects are independently sufficient for causation; he considers a hybrid account of causation where, *only jointly*, are the two sufficient.

Schaffer proposes that we combine some form of probability raising with some form of process connection. Specifically, Schaffer adheres to the general idea that, for a probabilistic account, a cause ought to raise the probability of its effect. Recall that, for most probabilistic approaches – like Lewis' – what we are required to do, in order to determine whether C causes E, is to compare the probability of E's occurrence when C occurs, to E's occurrence when C does not occur. If it is the case that in the closest possible worlds C's occurrence raises the probability that E will occur, we say that C causes E. But, as I have already intimated, Schaffer believes that a theory which is defined purely as probability raising is insufficient. He morphs the probability raising account into a hybrid by adding the necessity of a processes, which Schaffer calls a "connecting line" between cause and effect.

How we interpret this connecting line is up to us; Schaffer formulates his proposal in terms of an 'uninterpreted' analysis for which any number of different process accounts could be plugged in (even, perhaps, Hall's Intrinsicness thesis). The proposal, called PROPs, is as follows:

> "C causes E if and only if C is a (P)robability (R)aiser (O)f a (P)rocess for E (for short: C is a PROP for E)." (Schaffer 2001, 85)[35]

The idea is quite intuitive: in order that C be considered a cause for E, C must raise the probability of the process – the "E-line" – which is connected to E by actual events.

Schaffer introduces a procedure for implementing the PROP theory. We must start, he says, by determining our E-line. This line will be a process which includes E and all the connected events back to some earlier time. It will look something like this: (C^*, D_1, D_2, ... E). Schaffer then says that,

1) when the C event is the C* event (and thus part of the E-line), then C is a PROP for E if and only if "C is an essential part of the E-process: without C, if E still occurs it is via a different process entirely, rather than the same process slightly altered" (*ibid.,* 87).
2) When C is not part of the E-line, then it is a PROP for E, "if and only if C is a shield for the process: without C, the chance of that same process running to completion would have been less" (*ibid*).

So, to sum up, Schaffer's unified theory of causation proposes that C is a cause of E if and only if C is a probability raiser of a process for E. Specifically we are supposed to determine an E-line (some chain of events ending in E that goes back to

35 Schaffer goes on to offer his preferred account of PROPs, which the reader is invited to adopt, or discard, as she desires. The interpretation is: "C is a PROP for E if and only if (i) there is an extended event 'E-line' containing actual distinct events <C', D_1, D_2, ..., D_n, E> in pairwise nomic subsumption relations, (ii) there is an actual event C at time t_C, which is distinct from D_1, D_2, ..., D_n, and E (C may or may not be distinct from C'), (iii) ch(E-line)-at-t_C = p, and (iv) ~C□→ch(E-line) at $t_{(\sim C)} < p$." (Schaffer 2001, 85)

some earlier time) which may or may not include C. When the process does include C, then C is a PROP for E if and only if when C does not occur, E will not occur and if E does occur it will be by a different process entirely. When the E-line does not include C, then C is a cause of E if, and only if, when C does not occur the chance of the E-line occurring is less than when C occurs. In the latter of these two cases, C is called a *shield* for E.

I said our E-line had to contain the E event and "*all the connected events back to some earlier time*," because for Schaffer, all of the events must be actually occurring, distinct events. This may seem, as it did when I discussed Lewis, to eliminate the possibility of causation of, and by, omissions. Schaffer does not wish to make this strong an assumption and so proposes to deal with omissions as follows. When an event is an absence, we ought to "think of the E' incompatible with E as the *manifestation* of E's absence: if E is the absence of beer in the fridge, its manifestation is the fridge actually being stuffed with milk, grapes, and air" (Schaffer 2001, 86). This approach, so far as I can tell, does not get around those problems mentioned in §2.1 on Lewis: there will still be a great number of cases which will likely count as causation by, and of, omissions. This however, appears to be a conclusion which Schaffer swallows for he says "the only options with respect to absences are all or none" (2000a, 295, fn.12).

Because this 'all-or-none' approach drastically contradicts our everyday experience – (to use Schaffer's example) it seems natural to say that the gardener's failure to water my flowers caused them to die, while it seems entirely *unnatural* to say that the queen's failure to water the flowers caused them to die – Schaffer must offer some explanation. It comes in the form of pragmatism: "Since I never presumed that the queen would deign to water my flowers," he says, "to speak of this absence is to impart no information not already supposed" (2000a, 295).

The debate can go many directions form here, and it is hardly the place to trace them out in full. Nevertheless, I shall adopt this *manifestation* approach for the remainder of this essay, despite its possible shortcomings, leaving the issue for another day.

Schaffer's PROP theory, to recap, states that neither dependence nor process connection is independently sufficient for causation; it is only jointly that the two relations are sufficient for the causal relation. Thus, Hall's objections do not apply. For Hall, what it is about dependence that conflicts with the theses for production is the insistence that counterfactual dependence is sufficient for causation. But Schaffer recognizes that it is clearly *not* sufficient. What is sufficient for causation on the PROP theory is the combination of dependence with some process account; it is not enough to have counterfactual dependence among events, there must also be the connecting process: only as a pair are counterfactual probability raising and process connection sufficient for causation.

4.2. Analysis of Schaffer's Account

Let us see how this account might handle the preemption cases and the double prevention cases. For late preemption (Figure 3 – because it is more troubling) to see

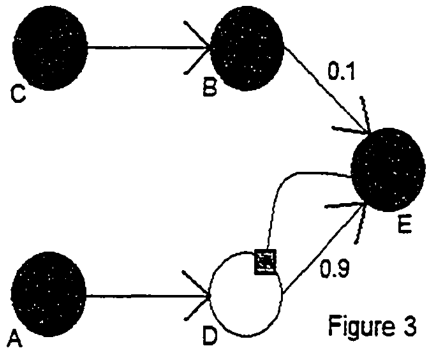

Figure 3

that C is a cause of E, we must first pick our E-line: let it be C→B→E. C is an essential part of the E-line, and so is a cause of E if and only if had C not occurred E would not have occurred, or E would have occurred by a different process entirely (not a slightly altered one). This is surely the case: had C not occurred, E would have occurred, but by a different process entirely (A→D→E). Thus, C is a cause of E. Moreover, the analysis avoids the conclusion that A (Phil's throw) is a cause of E, because had A not occurred, E would have occurred by the same process (C→B→E), except slightly altered (because with the absence of A, e.g., there will be no gravitational effects from Phil's rock to affect Bonny's rock).

Double prevention (Figure 1) is similarly dealt with; recall, Phil and Bonny are on a bombing mission. Keen-eyed Phil (C) spots fighter pilot Enemy (B), veers left and

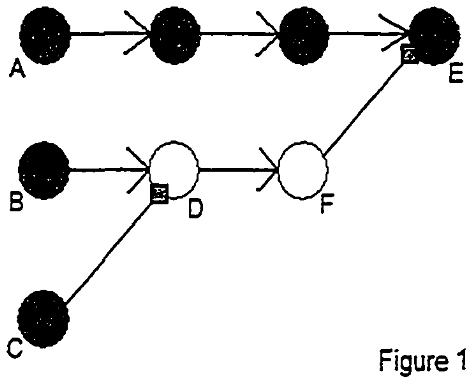

Figure 1

shoots him down, which prevents Enemy from continuing his flight path (D), and shooting down Bonny (F). Bonny (A) continues on her way and destroys the target (E). If Phil had not shot Enemy down, Enemy would have avoided Phil and shot down Bonny, and she would have thus failed to bomb the target.

Again we pick our E-line: all the events, including E, back to the time of A. The event in question, C, is not part of the E-line, so if it is a cause at all, it has to be a cause *qua* shield. Thus, for C to count as a cause of E, the chance of the E-line (A→E) running to completion must be less than what it would have been, had C not occurred; by assumption, if Phil had not shot down Enemy (say Enemy avoided Phil), then the chance of the target being destroyed by Bonny would have been less. So the approach seems to work.

Schaffer's PROP theory is subject to objections, however. The first objection, of which he is aware, brings back the case of early-preemptive double preventions. Recall the example: Bonny, a fighter pilot, is on a bombing mission. Phil and Wingman accompany her. Phil preempts Wingman in shooting down Enemy. Enemy almost certainly would have shot down Bonny, had neither Phil nor Wingman been present. Nevertheless, Bonny completes her bombing mission, due to Enemy's destruction. We

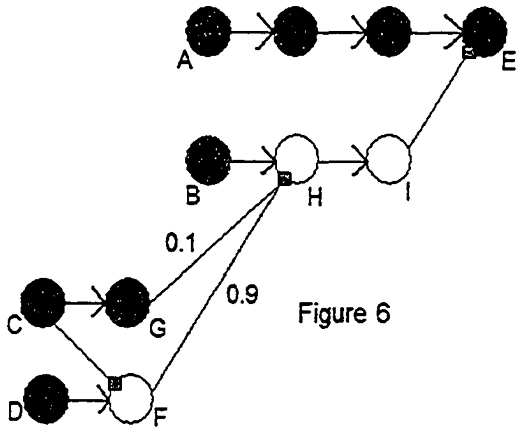

Figure 6

can make the situation probabilistic by stipulating that Phil is much less likely (0.1) to destroy Enemy than Wingman (0.9) because Phil is a worse shot. We get Figure 6.

The PROP theory does not get the intuitively correct answer in this situation (namely that C is a cause of E), because C, Phil, is a shield and thus extrinsic to the process A→E. That is, C, as a shield, must make it that the process A→E (Bonny's flight path and bombing) would be much less likely to occur without C; however, because D (Wingman) is far more likely to prevent H – his probability of success is higher after all! – C does not raise the probability of the process A→E: The process is *more* likely to occur if C had not occurred.

Schaffer's solution to the problem is not a new idea. He adopts the Lewisian approach of making the chain between C and E continuous. All we have to do is look for an event after the preemption, which depends on C and not the preempted alternative. In this case, it is event G. E, in turn, will counterfactually depend on G. He then further stipulates that we ought to be able to find, for any intermediary time T_Q some event Q which will depend counterfactually on the event preceding it, and upon which the event following it will depend counterfactually. In this instance Schaffer calls C a 'direct PROP' for E. By this definition, C will be a cause of E, because there is a continuous chain of direct PROPS from C to E. (This approach also makes the definition transitive.)

Adhering to the continuous modification, however, will not serve to solve two of the other pressing problems. One of these pressing problems is formed by changing the situation of Figure 6 to include the *late* preemption structure by stipulating that Phil fires a tiny bit earlier than Wingman; we get Figure 7. Here the original PROP theory will not work because D – which is causally idle – is an event which is far more likely to shield the E-line. Nor will the continuous amendment conclude that C is a cause of

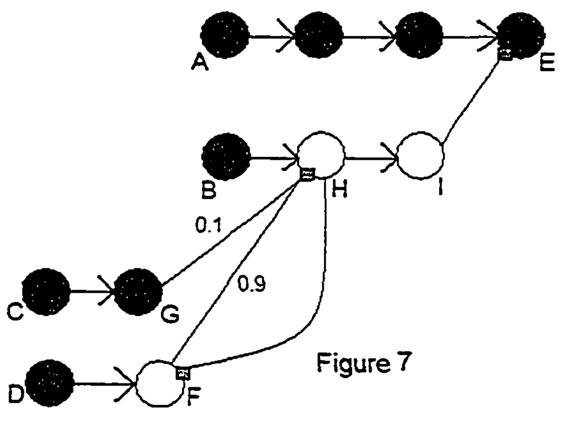

Figure 7

E because there is no event after the preemption (H, I, or F), upon which the prevention of H will depend; that is, because both the Phil-process and the Wingman-process are on the road to completion, and because it is only the destruction of Enemy by Phil's bullets that preempts Wingman's bullets, there will be no event in the Phil line which is late enough such that Enemy's destruction will depend upon it. Because of the (buried) late-preemption structure, E will depend on I, and I will depend on H, but H will depend on F, not G as is required. Thus, the continuous approach fails. Even if we were to make the case entirely deterministic, though the theory would not find that H depends on F, it would not be able to find dependence between H and G. (Notice here, that the reason the continuous approach fails for Schaffer is essentially the same as the reason it failed for Lewis; see § 2.1: p.22-3.)

The last case which poses a problem for the continuous modification, which I have not yet named (but which is a different flavor of the combination type problem), is again illustrated by taking Bonny and Phil as fighter pilots. This time, however, remove Wingman entirely, and add a bomb under Enemy's seat which would have gone off slightly later than the time at which Phil actually shoots down Enemy.[36] Again, stipulate that Phil is a poor shot, with a low chance of success (0.1), and also that the bomb under Enemy's seat is huge, and has a high chance of destroying Enemy (0.9). We get Figure 10 (where H is the bomb under Enemy's seat). Schaffer's PROP theory does not work here, again, because C is an extrinsic factor which can be affected by other elements in the causal situation (in this case it is the addition of H); because H is more likely to prevent the preventative chain B→D→F, the E-line does

[36] This variation is Hall's (2004b, 244-5); the added probabilities and the diagram are mine, however.

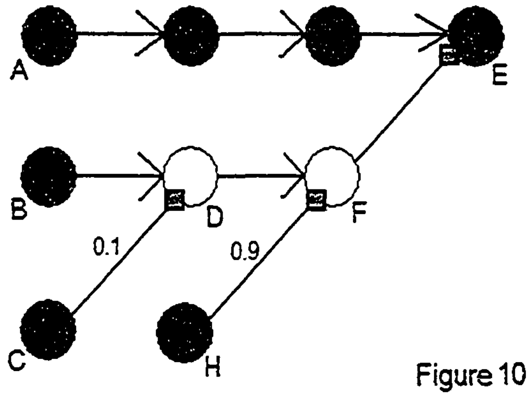

Figure 10

not depend on C. Moreover, because we have stipulated that H is a more likely preventative event than C, making the theory continuous will not escape the problem. Even though D depends on C, D is not a sufficient intermediary event because E does not depend upon it; and, given that F is more likely to be prevented by H than D, E will depend counterfactually on F, which will depend counterfactually on H (not D) as the definition requires. Thus, the continuous modification fails for Figure 10.

There is yet another problem that plagues the continuous solution as it occurs in both cases: our E-line changes dramatically. Let us assume that the approach works, and consider the cases, starting with early and late-preemptive double prevention cases (Figure 6 and 7). Presumably, these cases are similar enough to warrant that same E-line as in the original case (Figure 1), A→E. The change to a direct chain of PROPs however alters the E-line, in both Figures, to C→G→H→I→E. There is no sign of A (or the event that follow it) in the new E-line, and C has gone from being *extrinsic* to the process to an essential – intrinsic – part of it. The same can be said when we use the continuous approach for Figure 10. What was once a shield of the E-process A→E is now intrinsic to it (because the E-line would now be C→D→F→E). It seems clear that C is a cause, in all the cases, because it is a shield for the E-line; it is not what Hall would class as a 'producer' of E. But, adhering to the continuous approach makes C just that: a 'producer' of E. If we are to maintain the intuition that C as a shield (i.e, extrinsic to the process which causes E), then we must abandon the *continuous* approach.

Though Schaffer's account fairs considerably better all-around than Hall's, his theory's down fall – like Hall's – comes at the hands of examples which are the concatenation of the more primitive cases. Figure 7 is an explicit example of such a case; and, even Figure 10 is such a case. The relationship between the C event and the H event is a varied kind of preemption; C preempts H by preempting string of events before H is able to do so. Even thought here is no direct line of preemption between C and H, they nevertheless stand in a relationship of preemption.[37]

5. General Implications for Theories of Causation

What is left for prospects of spelling out a theory of causation? Hall, by my count, is worse off than Schaffer. Though Hall has offered useful insights into the nature of the five theses of causation, his conclusion that there are two concepts of causation runs into serious trouble when it is more fully spelled out. I think we ought

[37] For more on this, see Collins (2004)

to bear in mind Hall's criticisms while shying away from his proposed disunity of causation.

Schaffer's theory is in better shape than Hall's: at least his – Schaffer's – position is not self defeating. And yet, one half of Schaffer's theory – the shield definition – is subject to the same kind of criticisms which Hall raises against combining intrinsicness with dependence and omissions (clearest in Figure 10). H is, by all accounts, an event such that when it is present there is no dependence of E on C. This type of event is a variation of the 'extrinsic presence of enabling factors' type of event mentioned by Hall (see §3.1, and footnote 22). This, however, is called the problem of *extrinsic absence of disabling factors* for reasons that I have already given: if H were present, the dependence of E on C would be disabled.

So far as this discussion is concerned, any hopes of developing a theory of causation will have to come from the unified camp. And, given that we have a 'half-successful' theory (in the form of PROPs), the direction to take would be to modify the troublesome half of the definition of a shield, in an attempt to avoid the the effect that extrinsic factors (like the bomb under Enemy's seat) have upon the causal status of a shielding event.

The amendment to the shield definition that I support is one which is suggested by the continuous approach. Let us look back at the first case (Figure 7) again to see how we might draw out a solution which encompasses both cases. When we made the chain continuous in our first case of preemptive double prevention, what we said was that C was a PROP for H, and H a PROP for E; and, that C, in virtue of this chain, was

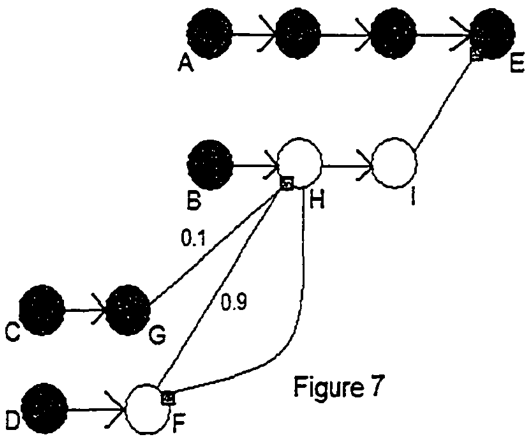

Figure 7

a PROP for E. However, if we look closer we can see that C is a PROP for H, not by being a shield for it, but by being an essential part of the process for it. That is, had C not occurred, H would not have occurred, and if it does occur, it would have been by a different process entirely (namely the process beginning at D). It is the event which begins the former process (C→G→H→I) obviously, which is the PROP, *qua* shield, for the original process A→E. Thus, we can say that the process A→E depends counterfactually on the shield C, which is an essential part of the process C→I. (Recall that because we have tried to maintain the intuition that C is not an intrinsic part of the process that produces E, we stop short of including E with the process C→I. Doing so allows us to hold that C is a cause of E *qua* a shield for the process A→E.)

It is easy to look at the case represented by Figure 7 to see why E depends on the event which starts the process that ends in I: it was Phil's actions and the events related to those actions, and not Wingman's actions or those related events, which prevented Enemy from preventing the bombing. My above formulation simply explicitly incorporates this thought.

The actual definition would be the same as Schaffer's with the emendation added to the shield definition. Thus,

C causes E if and only if C is a PROP for E;

And, we pick our E-line which contains events (C^*,D_1,D_2,... E). When:

1) the C event is the C* event, then C is an essential part of the process, and C is a cause of E iff, "without C, if E still occurs it is via a different process entirely, rather than the same process slightly altered" (2001, 87); and when,

2) the C event is not the C* event, then C is not an essential part of the E-line. Identify the process to which C is essential. Call this P_C. Then, C is a cause of E iff the chance of the E-line running to completion when P_C does not occur would be less, and if the chance is not less it is due to P_B, another process entirely (which is started by a completely different event, B). (Because C is essential to P_C, P_C will not occur when C does not occur.)

This new account, it should be remembered, is built upon the original definition of PROPs, and so does *not* require the *continuous* PROPs amendment which Schaffer suggests; the processes are to be understood in the normal way. (In fact, were it to require a continuous process in the manner that Schaffer posits, this amendment would not solve the problem which motivated it.)

The amendment to the definition of PROPs leaves intact the successes had by the theory for the cases of early and late preemption (because there is no change to when C is an essential part of the E-line). It, further, gives the proper conclusion for the

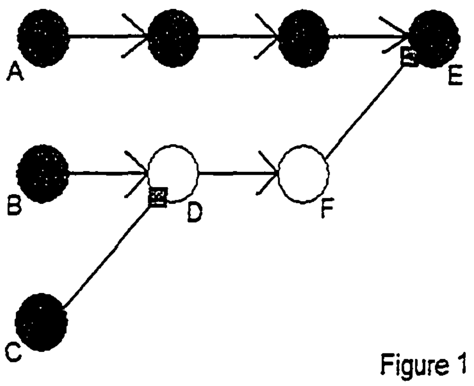

Figure 1

double prevention cases, while overcoming the problems of Figure 7 and Figure 10. Let's start with simple double prevention (Figure 1). Our E-line is, again, A→E. C is not a part of the E-line, so we identify the process to which C is essential: P_C is C→D→F. Now we see whether the chance of the E-line would have been less if P_C had not occurred. We can see that if the process P_C did not occur, then indeed the chance of the E-line running to completion would have been less.

Nor does the analysis of Figure 1 with my amended definition run into trouble when we add other factors which are (even more) external to the causal shield. If you like, let us return to the case of preemptive double prevention (Figure 7) by adding back in the presence of Wingman. We have the same E-line: A→E; P_C, lengthens slightly to include the event G becoming C→G→H→I. Even when we recall that Phil is very unlikely to shoot down Enemy (but he manages to anyways) and that Wingman

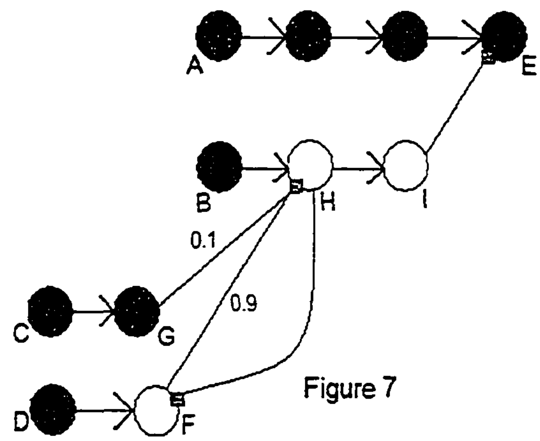

Figure 7

is very likely to shoot down Enemy, the analysis is not tripped up. The chances of the E-line running to completion without P_C would not have been less, but it would be because of another event D, which initiates an entirely different process P_D (which is D→F→H→I).

The same is true when we add the large bomb under Enemy's seat (Figure 10). The bomb, even though an external factor to Phil the causal shield, does not affect the causal status of Phil any longer. The E-line remains the same from Figure 1: A→E. It does not have less of a chance of running to completion when the process P_C, C→D→F, does not occur – but this is due to another event H, which is starts another process entirely P_H, (H→F).

This account also has the advantage that causation remains transitive. Indeed, as I have redefined it, even the case which prompted Schaffer to make his PROP theory continuous (i.e., the combination cases), are transitive. Consider: Phil's actions (C) will

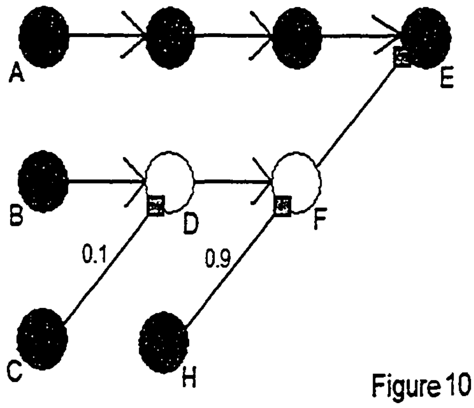

Figure 10

be a PROP for the destruction of Enemy (H), by being an essential part of the process which produces it (C→G→H); Enemy's destruction (H) will be a PROP, *qua* shield, for the bombing performed by Bonny (E), because without it, the E-line (A→E)would have less of a chance of running to completion, and the process to which H is essential would not have occurred (H→I). And, Phil's actions (C) are a PROP for the bombing (E), because had C not occurred, the process to which it is essential (C→G→H→I) would not have occurred, and the chance of the E-line running to completion would have been less. So, C causes H, H causes E, and C causes E; transitivity is maintained.[38]

Whether or not this account works, it can be understood as a further attempt at hybridization by implementing a form of stability in structure. As I understand the original PROP theory given by Schaffer, the aim was to take the successes of one

[38] I leave, as a piece of unfinished business, the claim raised in §3.1 that perhaps we do not want *all* causal cases to be transitive. I am of the mind that there are some cases which probably are *not* transitive. However, it is unclear, besides by using plain intuition, how we can mediate between the cases that we clearly want to count as transitive, and those that we wish to deny are transitive.

particular global theory and, while maintaining its successes, further them by stabilizing the definition against the addition of external factors to the causal line.[39] This is readily apparent in the first half of the PROP definition, when C is essential to the E-line. Whenever C is essential to the E-line, the cases is, in Hall's terms, one of production. With this in mind, we can, mentally, consider the diagrams which might constitute our blueprint structures, and then add elements to see what I mean. Consider the case where Bonny throws her rock, Phil is not present, and she breaks the window; we will have a specific structure much like a straight line with (at least) three neurons – the first being her throw, the last being the window breaking. The first neuron is essential to the structure of the diagram. Now, when we add some Phil-neurons to this structure of three Bonny-neurons, the causal status of C will not change because the same process, when Bonny was alone, remains. Moreover, if we were to massively disturb the stability of the diagrams by adding Phil and *removing* Bonny, we massively disrupt the processes present in the diagram (because we will have removed the three Bonny-neurons by removing Bonny), and at the same time, disrupt which actions we can correctly count as a cause of the breaking.

However, the original PROP account was still too unstable for all instances, as we have seen in the case of shields. Shields are necessarily extrinsic to a causal process; they cannot be an essential part of the 'producing' line for the effect. And, as

[39] I am fairly confident that Schaffer is stabilizing a global theory rather than broadening a local theory, because I do not believe, had Schaffer come into the discussion earlier (when he would have had to choose a purely local or purely global account), that he would have advocated a purely local account of causation. This instinct is due almost entirely to the fact that he was the first to present the double prevention cases (2000), and adamantly argue against the adequacy of local theories to handle the double preventions cases. He notes, in that same paper, the ease with which the global accounts deal with the cases.

Schaffer construes them in his original definition of PROPs, they are subject to serious objections like the extrinsic absence of disabling factors (found in Figure 10), which disrupt the stability of the structure. Picture how the structure is altered by the simple – and, for all intents, irrelevant – addition of the neuron H. Now, if we agree, in the case of double preventions, that Phil is a cause of the bombing, then it isn't too much of a stretch to get the same verdict for a different case of double prevention, where Phil is *absent*, but the bomb under Enemy' seat is present. (This structure is *exactly the same as* the structure of Figure 10, except that the C neuron removed entirely. Notice that it is still a case of double preventions, but the shielding neuron fires slightly later.) Thus, we seem to have two conflicting intuitions: when Phil is present he is a cause of the bombing; and, when the bomb under Enemy's seat is present, it is a cause of the bombing. And, not surprisingly, when we combine Phil and the bomb, our intuitions are muddled, because we have lost the stability in our structure.

By entwining the idea of a process within the definition of a shield, I attempt to regain the stability from the causal structure of Figure 1. Taking the case of double preventions where Phil is alone and there is no bomb under Enemy's seat, the entirety of Figure 1 is captured by our definition (i.e, we can account for the presence of all the neurons). This is the case, recall, because under my definition we take into account the process to which Phil is essential – thus we account for the events C, D, and F. Adding H to this, as we do in Figure 10, does not cause massive disruptions in stability because we have not added any other (significant) events or processes which interrupt the already present processes.

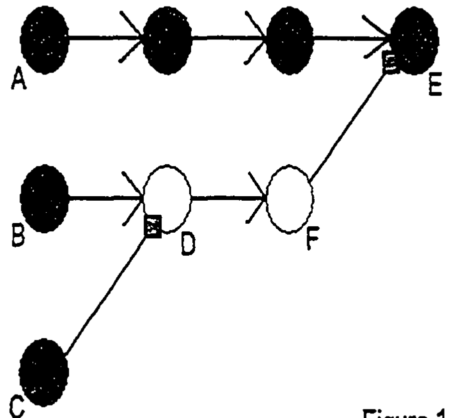

Figure 1

Speaking more generally, adding stability to causal structures is not a new idea. In fact, I have already suggested the idea of adding a certain level of stability across events. For example, in §2 (on Lewis) I briefly mentioned, since Lewis' original counterfactual account failed to give the intuitively correct verdict for some of the preemption cases, that he abandoned it for other pursuits, including the quasi-dependence account and the influence account. These were explicit attempts to stabilize the causal structures across events. (For further comments on this, see Hall [2004c].)

Now, I am unsure whether my amended account has served to 'stabilize' shields sufficiently enough to avoid further counterexamples. Clearly, it is *not* possible to fully stabilize a shield such that it is intrinsic to the causal process – this is what Schaffer's *continuous amendment* does, and it removes the shield from its shield status. If we are to maintain the shield as a shield, we must change the definition just enough such that the causal structure is stabilized so that further additions to the causal structure do not affect the status of the cause in question. So far as I can tell, my definition does just this – but it is a matter for creative minds to decide.

References

Collins, John. (2004). "Preemptive Prevention." in *Causation and Counterfactuals.* ed. J. Collins, N. Hall, and L.A. Paul. MIT Press, Cambridge, Massachusetts. 107-118

Dowe, Phil. (2000). *Physical Causation.* Cambridge University Press, Cambridge, UK.

Dowe, Phil. (2001). "A Counterfactual Theory of Prevention and 'Causation' by Omission." *Australasian Journal of Philosophy.* June; 79(2): 216-226

Fair, David. (1979). "Causation and the Flow of Energy." *Erkenntis* 14: 219-250.

Hall, Ned. (2002). "The Intrinsic Character of Causation." In Dean Zimmerman (ed.), *Oxford Studies in Metaphysics,* Vol. 1. Oxford: Oxford University Press. 255-300 .

Hall, Ned. (2004a). "Causation and the Price of Transitivity," in *Causation and Counterfactuals.* ed. J. Collins, N. Hall, and L.A. Paul. MIT Press, Cambridge, Massachusetts. 181-204

Hall, Ned. (2004b). "Two Concepts of Causation." in *Causation and Counterfactuals.* ed. J. Collins, N. Hall, and L.A. Paul. MIT Press, Cambridge, Massachusetts. 225-276

Hall, Ned. (2004c). "Rescued from the Rubbish Bin: Lewis on Causation." *Philosophy of Science.* Dec.; 71: 1107-1114.

Hitchcock, Christopher. (1995). "Salmon on Explanatory Relevance." *Philosophy of Science.* June; 62(2): 304-320

Hitchcock, Christopher. (2003). "Of Humean Bondage." *British Journal of the Philosophy of Science*; 54: 1-25.

Lewis, David. (1973). "Causation." *Journal of Philosophy* 70: 556-567. In Lewis, (1993). 193-204

Lewis, David. (1979). "Counterfactual Dependence and Time's Arrow." *Nous* 13: 455-467. Reprinted in Lewis (1986a): 32-66

Lewis, David. (1986a). *Philosophical Papers,* volume II. Oxford: Oxford University Press.

Lewis, David. (1986b). "Postscripts to 'Causation'." In Lewis (1986a): 172-213.

Lewis David. (1993). "Causation." *Causation,* eds. Sosa & Tooley. Oxford University Press, Oxford.

Lewis, David. (2000). "Causation as Influence." *Journal of Philosophy* 97: 182-197

Salmon, Wesley. (1984). *Scientific Explanation and the Causal Structure of the World.* Princeton, N.J.: Princeton University Press.

Schaffer, Jonathan. (2000a). "Causation by Disconnection." *Philosophy of Science.* June; 67(2): 285-300.

Schaffer, Jonathan. (2000b). "Trumping Preemption." *Journal of Philosophy* 97: 165-181.

Schaffer, Jonathan. (2001). "Causation as Probability Raisers of Processes." *The Journal of Philosophy* 98(2): 75-92.