DE FACTO DEPENDENCE*

Two friends throw rocks at a window; Suzy's rock hits the window, while Billy's sails harmlessly through the now empty frame. It was Suzy's throw, not Billy's, that caused the window to break. But counterfactually, the two throws seem on a par. Had neither occurred, the window would not have broken; had either occurred without the other, it (still) would have. This is an instance of the *preemption problem* for dependence-style accounts of causation. Suzy's throw is a cause despite being no more depended on than Billy's.

But is it really so that Suzy's throw is no more depended on than Billy's? Take away Billy's throw while holding other things fixed, and the window still breaks. But if we take away Suzy's throw while holding other things fixed, it does not break. This is because one of the things we have to hold fixed is that Billy's rock never touches the window. Had Suzy not thrown, with Billy's rock never making contact, the window would have had nothing to fear from either rock, so it would not have broken.

Effects do not always counterfactually depend on their causes. But they do (seem (so far) to) *de facto* depend on them—to depend on them with the right things held fixed. Of course, a story will be needed about what the "right things" are. That is the business of the next few sections, after which I compare the de facto theory to some other recent responses to preemption.

I. ANALOGY

Preemption is a problem for the counterfactual theory of causation, familiar from the work of David Lewis.[1] Compare the problem facing another counterfactual theory, also discussed by Lewis. One starts out thinking that

> *x* is disposed at time *t* to give response *r* to stimulus *s* if and only if, if *x* were to undergo stimulus *s* at time *t*, *x* would give response *r*.[2]

[1] "Causation," this JOURNAL, LXX, 17 (October 11, 1973): 556-67, reprinted with "Postscripts" in *Philosophical Papers, Volume 2* (New York: Oxford, 1986), pp. 241-69. See also Ardon Lyon, "Causality," *British Journal for the Philosophy of Science*, XVIII (1967): 1-20.
[2] Lewis, "Finkish Dispositions," *Philosophical Quarterly*, XLVII (1997): 143-58. See also Charles B. Martin, "Dispositions and Conditionals," *Philosophical Quarterly*, XLIV

But, Lewis notes, "stimulus *s*...might chance to be the very thing that would cause the disposition...to go away. If it went away quickly enough, it would not be manifested."[3] This is the *finkishness problem*, so called because the theory overlooks dispositions that "fink out" when the stimulus is applied.[4] An example is Saul Kripke's[5] killer yellow, which destroys the brains of those who set eyes on it before any visual experiences can occur.

Finkishness and preemption have a great deal in common.[6] Both make trouble for analyses of the same basic form: *x* is *F* if and only if $(Ax > Bx)$. Both arise because *x* is *F* in virtue of some underlying *G*, yet *x* would not have been *G*, had it been *A*. Where the problems differ is in the identity of *F* and *G*. Finkishness has *F* = *x*'s disposition and *G* = its grounds; preemption has *F* = *x*'s causing of *y* and *G* = ...well, that is what we have to explain.

Socrates distinguishes between "the cause" of an effect, and "that without which the cause would not be a cause" (*Phaedo* 98e). Take the fact that Billy's rock never touches the window. This is not a cause of the window's shattering; it does not contribute to the shattering in any way. But it does help to make something else the cause. It is a cause *maker* (read like "king maker").[7] The no-show status of Billy's rock helps to make Suzy's throw the cause by putting the effect in need of it. So, proposal: where finkishness has

> *F* = *x*'s dispositional property
> *G* = the disposition's grounds, that without which *s* would not be sufficient for *r*

preemption has

> *F* = *x*'s causal relation to *y*
> *G* = the relation's grounds, that without which *x* would not be required by *y*

Preemption happens because to take away the cause is, sometimes, to take away more. It is to take away one of the reasons it *is* a cause,

---

(1994): 1-8; and Robert K. Shope, "The Conditional Fallacy in Contemporary Philosophy," this JOURNAL, LXXV, 8 (August 1978): 397-413.

[3] "Finkish Dispositions," p. 144.

[4] There is a parallel problem, Lewis notes, about finkish undispositions: previously undisposed *x* might acquire from *s* the disposition to respond to *s* with *r*.

[5] This example is from unpublished lectures on secondary qualities.

[6] They are usefully juxtaposed in John Collins, "Preemptive Preemption," this JOURNAL, XCVII, 4 (April 2000): 223-34.

[7] Cause makers are called "ennoblers" in my "Advertisement for a Sketch of an Outline of a Proto-Theory of Causation," to appear in Collins, Ned Hall, and Laurie Paul, eds. *Causation and Conditionals* (Cambridge: MIT, forthcoming).

namely, that other would-be causes miss the mark, leaving *y* with no other source for what *x* provides. The effect cannot be expected to follow *x* out of existence, if the reasons for its dependence on *x* go out of existence first!

The claim so far is that preemption is *analogous* to finkishness. One might try to turn the analogy into an identity. What is there to *x*'s causing of *y*, beyond the fact that *x* and *y* both occur? A *disposition* on the part of the circumstances to lose *y* if "stimulated" by the loss of *x*.[8] For *x* to cause *y* is for it to engage with that disposition. Whatever grounds the disposition contributes thereby to *x*'s status as cause.

II. HOLDING FIXED

No one is happy with existing theories of preemption. But about finkishness there is a certain amount of optimism. The form of the solution is thought to be this: *x* is disposed to produce *r* in response to stimulus *s* if and only if

> Were *x* to be presented with *s* while retaining grounding property *G*, then *x* by virtue of being *G* would give response *r*.[9]

Now, though, one must specify *G*. One does not want to do it in terms that presuppose the notion of a disposition, although if the presupposing was minor and indirect, we might still be interested. Best would be a definition that did not appeal to dispositionhood at all. Let us review some of the options.

(1) $G = x$'s *total intrinsic nature*. This holds too much fixed. Some dispositions cannot be manifested without their bearers changing intrinsically. (Lewis gives the example of elasticity.) It also holds too little fixed, for some dispositions are grounded in extrinsic properties. An example is weight as conceived by operationalists: the disposition to register thus and so many pounds when put on a scale. This must mean on a *local* scale, since things weigh less on the moon than on Earth. But perhaps my locality would change if someone tried to weigh me. (You would have to chase me to the moon to get me on to that scale.) The disposition is there, but because it is grounded in external circumstances that change when the stimulus is applied, the counterfactual fails.[10]

(2) $G = x$'s *total categorical nature*. This again holds too much fixed. Some dispositions are manifested through categorical changes in

---

[8] Why not also a disposition to "gain" *y* if stimulated by the addition of *x*? This idea has some merit, I think, but let us stick for now with the negative disposition.

[9] Lewis, "Finkish Dispositions"; Collins, "Preemptive Preemption."

[10] Extrinsic dispositions are further discussed in my "Intrinsicness," *Philosophical Topics*, XXVI (1999): 479-504; see especially pp. 493-95.

their bearer: changes of shape and size, for instance, in the case of elasticity. The proposal also holds too little fixed, since some dispositions are grounded in hypothetical properties. Glass is fragile because of the way its chemical bonds are disposed to break under pressure. Warfarin is poisonous because of its anticoagulant powers.

(3) G = *any property of x such that were x to undergo s while retaining G, it would give response r.* No, because any x that is not undergoing s has a property like that: the material conditional property of producing r if undergoing s. A match is not disposed to bark like a dog on being struck. But that is what it would do, if we held fixed its property of either barking like a dog or not being struck.

(4) G = *any natural enough property of x such that were x to undergo s while retaining G, it would give response r.* The problem is not just that 'natural enough' is so obscure. It is that we get false positives even with quite natural Gs. Suppose this soft-boiled egg had been dropped, with its actual ovoid shape held fixed. The energy of the fall would have gone into the floor; the linoleum would have sagged and sprung back, flinging the egg into the air. The egg would have bounced even though it was disposed to break.

Option (4) yields not enough clear verdicts, and too many clearly wrong verdicts. Should we conclude that naturalness is no help? That would be too quick. Suppose that the Gs eligible to be held fixed could be identified (never mind how) using *comparative* naturalness rather than *absolute*. This would help with the first problem because whether G is more natural than H is a clearer matter than whether G is simply natural. It would help with the second problem because even a quite natural property can be *less* natural than some competitor. Our problems would be solved, then, if Gs arguing in favor of a dispositional link between s and r could be played off against Hs arguing against such a link, with the more natural condition prevailing. I do not know how to do this for finkish dispositions in general, but when it comes to the special case of preemption, a strategy does suggest itself.[11]

---

[11] I propose to postpone (= ignore) the question of what is the best thing to mean by 'natural'. This is partly because I am uncertain about it, in particular about the extent to which cognitive and cultural factors are allowed to come in. I get the term 'natural' and the idea that naturalness has a place in the metaphysician's toolkit, from Lewis—see his "New Work for a Theory of Universals," *Australasian Journal of Philosophy*, LXI (1983): 343-77, and (with Rae Langton) "Defining 'Intrinsicness'," *Philosophy and Phenomenological Research*, LVIII (1998): 333-45.

III. MAKE OR BREAK

Causation involves a disposition on the part of the circumstances to respond with the loss of e to a stimulus consisting in the loss of c. One identifies circumstances as so disposed by playing cause makers G off against Hs arguing for an opposite verdict: cause *breakers*, let us call them.[12] G is a cause maker if and only if (a) it actually obtains (G is a fact), and (b) had it been that $\neg Oc$ & G, it would have been that $\neg Oe$ (for short, $\neg Oc >_G \neg Oe$[13]). One can see how it would create a presumption in favor of c's causing e that e depends on it modulo a factual condition G. What H might do to undermine the presumption, however, is quite unclear. I propose to sneak up on the answer slowly.

Normally in thinking about causation, we conceive of c as supplying something that e has need of, and would not otherwise get. Preemption makes matters a little more complicated, but only a little. Other events send the needed something out along alternative routes. Holding G fixed blocks those routes, so we are returned to the previous case: c makes an essential contribution, without which e would not have occurred.

But although that is how it is supposed to work, the requirement of dependence modulo G supports an almost opposite scenario. The effect was going to happen anyway, when c came along to threaten it. It did this by creating a circumstance G that, left unchecked, would have sidetracked the process that was leading to e. Of course, putting the effect in jeopardy is not all that c did, or it would not even resemble a cause. It also "saved" e by taking it out of the jeopardy. A c that threatens e with one hand (its G hand) while saving it with the other is certainly not a cause of e; but the effect does depend on it holding G fixed. This is illustrated by an example of Hartry Field's.[14]

> "Bomb": Billy plants a bomb under Suzy's chair; Suzy notices it and moves out of range before it has a chance to explode; a few hours later, she goes to a prearranged medical checkup, and is found to be healthy.

Let e be the event of the doctor's issuing a glowing report on Suzy's state of health. The glowing report is not caused by Billy's planting of the bomb. But it is in a certain (strained) sense *contingent* on Billy's

---

[12] They are called "enfeeblers" in "Advertisement for a Sketch of an Outline of a Proto-Theory of Causation."

[13] '>' is the counterfactual conditional operator; '$A > B$' says that, if it had not been that A, it would not have been that B. '$A >_Y B$' is in general short for '$(A \& Y) > B$'.

[14] The example dates back to a seminar that Field gave at CUNY Graduate Center in the early 1990s.

action. If we hold fixed the fact that her chair exploded, Suzy would have been injured had she not been tipped off by the bomb. No doubt this is convoluted reasoning—one hardly wants to thank the bomb for alerting Suzy to a threat that it itself poses—but we have yet to see where it goes wrong.[15]

Here is a proposal, stated first in metaphorical terms and then gradually literalized. The problem with Billy's planting of the bomb is that it leaves us just as far from the effect as we would have been without it; "the road taken" is just as long as "the road not taken"—indeed, the one road includes the other. A little less metaphorically: everything the effect *would* have needed, had the bomb not been planted, it *continues* to need with it planted. One can cash the metaphor further by treating needs as events depended on under particular stated conditions.[16] G and H having been used for obtaining conditions (facts), let us introduce K for conditions that may or may not obtain.

> K puts e in need of x if and only if $\neg Ox >_K \neg Oe$, that is, e would not have occurred, had x not occurred in K-type circumstances.

If F is the (nonobtaining) condition of c failing to occur, and H is an appropriately chosen obtaining condition, then

> 'what the effect would have needed' = the events that F puts e in need of, and 'what the effect does need' = the events that H puts e in need of.

To say that 'Everything e would have needed, it continues to need' is to say that the second set subsumes the first, and with no help from c. What this means at an intuitive level is that the need for c is over and above what would, but for c, have been *all* the effect's needs. A need that is piled arbitrarily on top of what would otherwise have been all the needs is trumped up or artificial. The objection to c is thus very simple: it meets no real (nonartificial) need.

Now for some refinements and qualifications. History has a branching time structure, we assume. There is the trajectory actually taken through logical space; branching off of that are various alternative trajectories corresponding to other ways things could have developed. One branch in particular corresponds to the way things would

---

[15] Hall—"Causation and the Price of Transitivity," this JOURNAL, XCVII, 4 (April 2000): 198-222—offers a detailed discussion of this sort of case.

[16] If it seems odd to think of events as needs, remember that 'need' can mean thing that is needed. ("The dogsled was piled high with our winter needs.") Needs in the ordinary sense do not exist in our system. Their work is done by events considered under counterpart relation C, the relation of meeting-the-same-need-as.

have developed if c had not occurred. This, after it begins to branch off from actuality, is called the *fallback scenario*; the *actual scenario* is what happens after the branch-off point in actual fact. The events e depends on in the fallback scenario are the *fallback needs*:

$$\text{FAN} = \{x \mid \neg Ox >_F \neg Oe\}$$

(remember that F says that c fails to occur).[17] The effect's *actual needs*, for a given choice of H, are the actual-scenario events that it depends on modulo H:

$$\text{HAN} = \{x \mid \neg Ox >_H \neg Oe\}[18]$$

H makes the need for c artificial if and only if HAN covers FAN with c to spare; or, what comes to the same, FAN coincides with a c-free subset of HAN. The point either way is that the effect's actual needs c *apart* subsume its fallback needs. The would-be cause "makes itself indispensable" in the pejorative sense of puffing the effect's needs up beyond necessity, until finally it too can claim to have played a role.

So, what does the glowing medical report depend on, if Billy does not plant the bomb? Some pills have to be swallowed; Suzy must dodge the falling piano; the doctor's alarm clock has to go off; and so on. The problem with Billy's planting of the bomb is that it does not diminish these needs one iota. One does not even have to hold anything fixed; HAN contains all the fallback needs even if H is the null fact (the fact that always trivially obtains).[19] It also contains them on many other choices of H, including, ironically, the fact that put e in need of Billy's action in the first place: that Suzy's chair explodes. There are plenty of facts exposing the need for c as artificial.

IV. COUNTERPARTS

I said that the fallback needs (FAN) should *coincide* with a c-free subset of the actual needs (HAN). This has to be understood in a particular way. Suppose that Suzy gets sick if she does not stay hydrated. She sets her Palm Pilot to remind her to take H_2O at noon. Had Billy not planted the bomb, Suzy would have been in the chair

---

[17] It might be wondered whether FAN contains the events it is supposed to: the events that e would have depended on had c not occurred. x is in FAN if and only if $(\neg Ox \& \neg Oc) > \neg Oe$, while it is $\neg Oc > (\neg Ox > \neg Oe)$ that defines would-be dependence on x. The export-import law for counterfactuals—$P > (Q > R)$ if and only if $(P \& Q) > R$—brings the two requirements into alignment. I assume that the law is close enough to correct for our purposes, or anyway that the indicated consequence is close enough to correct.

[18] Remember that FAN and HAN are limited to events occurring *after* the branch point (the point at which the nearest c-less world peels off from actuality).

[19] It simplifies matters to pretend that there is only one such fact.

at noon; the need would have been met by an act of water drinking, water being the one hydrous stuff available in the room. Since Billy did plant the bomb, Suzy was out on the sidewalk when the need arose; the need was met by eating Italian ice, that being the one hydrous stuff available outside. Assuming that eatings are essentially not drinkings, the event actually needed is not the one that would have been needed, had Billy not planted the bomb. But then e has fallback needs distinct from all of its actual needs, whence it would seem that that HAN does not cover FAN at all, let alone with c to spare.

I answer that different *event* does not have to mean different *need*. Suzy's water-drinking and ice-eating meet the same need, as that need manifests itself in their respective scenarios.[20] The need for c is artificial if for each x needed in the fallback scenario, the effect's actual needs include, not perhaps that very event, but *an* event meeting the same need. Rather than each x in FAN having to be *identical* to a member of HAN other than c, each such x needs a *counterpart* in HAN, where counterparts are events meeting the same need.[21]

What does it mean for events to meet the same need? The idea is this. Fallback needs can be paired off with actual ones in ways that preserve salient features of the case: energy expended, distance traveled, time taken, place in the larger structure of needs. One wants to preserve as many of these features as possible, while finding matches for the largest number of fallback needs. One asks: How much of the fallback structure is embeddable in this way into the actual one? The maximal embedding determines which events meet the same need. (Should the actual structure succeed, with no help from c, in absorbing *all* of the fallback structure, the need for c is artificial.)

---

[20] Also, same event does not have to mean same need. An event that would have met one need in c's absence might in the actual scenario meet another (or it might meet no need at all). It is better not to treat needs as separately existing entities, but suppose for a moment that we do: z_w is the need met by event z in world w. Then what is claimed in the text is that x_u can be y_v even if $x \neq y$, and the claim of this note is that x_u can fail to be y_v even if $x = y$.

[21] I shall be taking counterparthood to be symmetric and one-one. But there might be reasons for relaxing these requirements. Take first symmetry. There might be an x in FAN whose closest actual correspondent meets, not the same need as x, but an even larger need: a need with the need met by x as a part. This closest actual correspondent ought to qualify as a counterpart of x. So, the argument goes, counterparts should be events meeting *at least* the same need, which makes counterparthood asymmetric. The one-one requirement seems overstrict for a similar reason. It might take a pair of events to meet the need x meets all by itself in the fallback scenario; or vice versa. I propose to ignore these complexities.

V. THE DE FACTO THEORY

I have been downplaying the fact that artificiality is relative to a choice of *H*. But the relativity is important to our larger strategy. That *e* is put in need of *c* by an obtaining condition *G* testifies to the existence of a causal relation between *c* and *e*. Sometimes, though, the testimony misleads. What distinguishes the *G*s that do signal a causal relation from the ones that do not (the "deceptive" *G*s)? An idea that did *not* work was to credit all *G*s that were sufficiently natural. But a related idea held promise. That the need *c* meets is artificial by the lights of an obtaining condition *H* testifies to the *non*existence of a causal relation. Rather than ask *G* to be natural in itself, let us require it to be more natural than the *H*s, if any, testifying the other way.

So: one event *de facto depends* on another if and only if it is put in need of the other by a *G* more natural than *H*'s (if any) exposing the need as artificial. That can be taken as a definition. The proposal is that

(DF) *c* is a *cause* of *e* if and only if *e* de facto depends on *c*.[22]

Later sections will compare (DF) to other recent theories of causation. Here I try it out on two well-known traditional challenges, *preemption* and *overdetermination*.

An example of preemption is given in our opening paragraph. How is it that Suzy's throw caused the window to break, when Billy threw only a moment later? The fact is that Billy's rock never comes into contact with the window; it never even gets close. That Billy's rock never gets close puts the effect in need of Suzy's throw. But is the need perhaps artificial, that is, superadded to what would otherwise have been all the effect's needs?

The effect's needs if Suzy had not thrown are for Billy's throw, his rock traveling toward the window, its hitting the glass, and so on. These would seem to recur in the actual situation as needs for Suzy's throw, her rock's traveling toward the window, and so on. But then Suzy's throw meets the same need as was met in the fallback scenario by Billy's throw. The need was thus preexisting; it would still have had to be met even if Suzy had elected not to throw. If someone thinks that the need met in the fallback scenario by Billy's throw does *not* recur in the actual situation, that is fine, too. Suzy's throw cannot be over and above the fallback needs unless those needs are still present; and on the suggested hypothesis, one fallback need, at least, is eliminated.

---

[22] I leave the usual qualifications to be understood; both events should occur, for instance, and each should be possible without the other.

This sort of pattern arises often enough to call for some new terminology. Say that c *meets* a fallback need if and only if it meets the same need as some x in FAN would have met in its absence; and that it *cancels* a fallback need if and only if there is an x in FAN such that *no* actual event meets the same need as x would have met in c's absence. It is easy to check that the need for c is artificial if and only if c does not address any fallback needs, meaning that it neither meets any fallback needs nor cancels any. Now, suppose it is known that c meets a certain fallback need if any actual event does. Then we can conclude at once that the need for c is not artificial. For if c does not meet the need, then nothing meets it; and if nothing meets it, then (given that the effect does occur), the need was cancelled.

Now we turn to the second traditional challenge. Overdetermination occurs when an effect depends on two events taken together, but not on either alone; and (what distinguishes it from preemption) neither of the two events can lay claim to being more of a cause than the other.

It might happen, for example, that Billy decides to throw a little earlier, so that his rock hits the window at the very moment that Suzy's does. Billy's throw is not a cause by the lights of the counter-factual theory, for the effect does not depend on it (we can suppose there are no chains of dependence, either). Suzy's throw is similarly disqualified. But then what does cause the window to break? Not the conjunction of the two throws, since the effect could too easily have occurred without it. Not the disjunction, because we are hard put to regard the disjunction as a genuine event.[23] Could it be that nothing causes the window to break? This goes somewhat against the grain. An event that was caused (the breaking was not a miracle!) should, one feels, have causes. I do not mean to be blaming our difficulties here on any particular theory; I agree with Lewis that the case is *intuitively* undecidable. The question is what to make of this datum. Lewis contends that overdetermination can be left as "spoils to the victor." But why not aim for a theory that shows a similar indecision to our own?

(DF) calls Suzy's throw a cause if and only if some G putting the effect in need of it is more natural than any H making the need artificial (likewise for Billy's throw). It is not hard to find suitable Gs. The shattering depends on Suzy's throw modulo the fact G_s that Billy's rock does not hit the window unaccompanied, and on Billy's

---

[23] Lewis, "Events," in his *Philosophical Papers, Volume 2*, pp. 241-69; see especially section VIII, "Events Are Not Disjunctive."

throw modulo the fact G_b that Suzy's rock does not hit unaccompa-
nied. It is not hard to find suitable Hs either; indeed, we have already
found them. G_s makes the need for Billy's throw artificial, and G_b does
the same for Suzy's throw. For suppose that Billy had not thrown. The
shattering would have depended on Suzy's throw, the forward motion
of her rock, and the like. Holding fixed that Billy's rock does not hit
alone (that is G_s), these events are still needed in the actual situation.
Thus $G_s = H_b$, and for similar reasons, $G_b = H_s$.

Assuming that these are the most natural cause makers (breakers)
to be had, is the effect put in need of Suzy's throw by a fact more
natural than any fact making the need artificial?

That depends. One reading of 'more natural' is *strictly* more natu-
ral. If that is what is meant, then neither throw is a cause; each prima
facie connection is broken by a fact exactly as natural as the one that
established it. But the phrase could also be taken weakly, to mean 'at
least as natural as'. It is true that each throw occurs under conditions
modulo which the effect ignores it. But then each also occurs under
conditions no less natural modulo which the effect needs it. Tie goes
to the runner on the weak reading, so we have two bona fide causes.
Our intuitive indecision about whether to treat overdeterminers as
causes reflects semantic indecision about what to mean by 'more'.[24]

VI. THE COUNTERFACTUAL THEORY

Event e depends counterfactually on event c if and only if it would not
have happened, had c not happened.[25] If we write $R*$ for the ances-
tral of a relation R—the relation x bears to y if and only if they are
connected by an R-chain—then the counterfactual theory says that

(CF) c is a cause of e if and only if e depends* on c.

It is understood that c and e both occur, that they are suitably
distinct, and that various unnamed other conditions are met. Before
considering some differences between (CF) and (DF), I want to point
out a respect in which they agree. If e depends on c (and the
unnamed other conditions are met), then c is a cause of e according
to both theories.

This is clear on the counterfactual theory because a relation entails
its ancestral. But why does dependence entail causation on the de
facto theory? If e depends on c, then the null fact (the fact that always
trivially obtains) puts e in need of c. It now has to be shown that no

---

[24] This is intended more as a rational reconstruction of our intuitive response
than a psychological explanation of it.
[25] "Dependence" is counterfactual unless otherwise noted.

comparably natural H exposes the need as artificial. But no H can expose the need as artificial. Artificiality presupposes fallback needs—events that e depends on in c's absence—and e does not even occur in c's absence.[26]

Now let me list some features of the de facto theory that, depending on one's other commitments, might seem to recommend it.

Not all preemptive causes register with the counterfactual theory. Early or "cutting" preempters get through, but late preempters, as in our opening example, are reckoned ineffective. A dependence chain from Suzy's throw to the shattering cannot be found, because there is no point whereafter only her rock is positioned to break the window. I am not claiming that the de facto theory catches all preempters of whatever type; I have not seen all the types. It does, however, hold out hope of catching them all.

Second, (DF) is a "sine qua non" account; the cause is a that without which not. (CF) gives up on the sine qua non idea because of preemption. But this is to snub the intuitions that initially recommended the theory. As Judea Pearl says: "...why should the existence of a counterfactual dependence chain be taken as a defining test for a concept as crucial as 'actual cause', by which we decide the guilt or innocence of defendants in a court of law? The basic counterfactual criterion does embody a pragmatic rationale;...we would like to encourage people to watch for circumstances where their actions could make a substantial difference. Once the counterfactual dependence between the action and the consequence is destroyed..., what good is it to insist on intermediate counterfactual dependencies along a chain that connects them" (*op. cit.*, p. 313)? De facto-ists think that ancestralizing is an overreaction. Preemption can be dealt with another way.

The counterfactual theory makes causation out to be transitive. But recent years have seen a number of counterexamples to transitivity. Consider again "Bomb." Billy's putting a bomb under Suzy's chair was a cause of her moving away; her moving away was a cause of her later being pronounced healthy; but his putting the bomb under her chair was not a cause of her later being pronounced healthy. The de facto theory can take these examples at face value. That e de facto depends

---

[26] Suppose that H does somehow manage to make the need artificial. That upsets the de facto dependence of e on c only if H is as natural as the null fact. Since dependence modulo the null fact entails dependence modulo every fact independent of c—every fact G such that $\neg Oc > G$— all such facts would have to be less natural than H before the de facto dependence claim was affected.

on a de facto dependent of c leaves the question of e's de facto dependence on c wide open.[27]

Even if some form of dependence is sufficient for being *a* cause, it is not sufficient for being *the* cause. An event c is not *the* cause of e unless it is proportional or commensurate with e. (This might be taken to mean that c is not screened off from e by its determinables, while its determinates are screened off from e by c itself.[28]) It is unclear, though, how the proportionality constraint can apply when e depends on c only indirectly. There are likely to be lots of dependency chains, assigning c lots of different successors. With which of these successors does c have to be proportional if it is to come out proportional to e? The de facto theory avoids this problem by insisting on one-off dependence.

VII. THE POTENTIAL ANCESTOR THEORY

One strategy against preemption (mine) is to bring hidden dependence relations to the surface by holding certain things fixed. A different strategy tries to bring them to the surface by holding some things fixed while changing others.[29] What changes is that other candidate causes are taken away. What is held fixed is that no "new" events—none, anyway, on which the effect depends*—are allowed to occur. This works well for the broken-window example. Had Billy not thrown, the effect would have depended* only on Suzy's throw and other actual events. Had Suzy not thrown, however, the effect would have depended* on events that did not occur, like the window's being hit by Billy's rock.

But there seem to be other examples where the preempted "cause" would not need new events to generate e. Suppose we are playing the kind of card game where success depends on quick reactions. Suzy and Billy are on one team, you and I are on the other. The Queen of Hearts has just turned up, and the winning team is the one that touches it first. Suzy and Billy both lunge at the card, with Suzy's finger arriving a moment earlier. It was Suzy's lunge that ended the

---

[27] There is no reasonable compounding operation on facts with the property that, if d depends modulo Q_1 on c, and e depends modulo Q_2 on d, then e may be expected to depend modulo $Q_2 \bullet Q_1$ on c. Any reasonable compounding operation has $Q \bullet Q = Q$, so it is enough to show that dependence modulo Q is not always transitive. But we know it is not when Q is the null fact; this is just the ordinary intransitivity of counterfactuals. For more on transitivity, see Michael McDermott, "Redundant Causation," *British Journal for the Philosophy of Science*, XL (1995): 523-44; and Hall, "Causation and the Price of Transitivity."

[28] See my "Cause and Essence," *Synthese*, XCIII (1992): 403-49, and "Mental Causation," *Philosophical Review*, CI (1992): 245-80

[29] Jonardon Ganeri, Paul Noordhof, and Murali Ramachandran, "For a (Revised) PCA-analysis," *Analysis*, LVIII (1998): 45-47

game, causing us to pound the table in frustration. If Suzy's lunge had not occurred, though, we would still have pounded the table as a result of events all of which actually occurred, prominent among these being Billy's lunge.

A second example is due to Alex Byrne and Ned Hall.[30] Suppose that c-type events produce e-type events directly, that is, with no help from intermediaries. The same is true of c'-type events, but it takes slightly longer: the effect follows in 1.1 seconds rather than 1 second exactly. If c and c' take place at noon and e occurs precisely 1 second later, then (other things equal), c is the preempting cause and c' is the preempted. But had c not occurred, e still would have, and without any dependence on nonactual events.

What does (DF) say? Consider first the card game. An obvious choice of G is the fact that over the whole course of the game, Billy never touches the Queen of Hearts. Holding that fixed, if not for Suzy there would have been no table pounding. In the Byrne/Hall example, we can let G be the fact that e does not occur more than a second after c'. Had c failed to occur in G-type circumstances, e would not have followed, for the slower-acting c' would not have had time to produce it. Unless the need can be made to look self-inflicted by an equally natural H (I cannot think of one), e de facto depends on c and so is one of its effects.

VIII. THE HASTENER THEORY

A third approach defines causes as events in whose absence e either would not have occurred at all, or would not have occurred so early.[31] If we take "never occurring" to be the limiting case of delay, the theory defines causes as events without which the effect would have been delayed. Since Suzy's lunge and faster-acting c make their respective effects happen earlier than they otherwise would have, the hastener theory can deal with the examples of the last section. But neither it nor the potential-ancestor theory can deal with Jonathan Shaffer's[32] trumping preemption:

> Imagine that it is a law of magic that the first spell cast on a given day [matches] the enchantment that midnight. Suppose that at noon Merlin casts a spell (the first that day) to turn the prince into a frog, that at 6.00pm Morgana casts a spell (the only other that day) to turn the prince into a frog, and that at midnight the prince becomes a frog. Clearly, Merlin's spell...is a cause of the prince's becoming a frog and Morgana's

[30] "Against the PCA-analysis," *Analysis*, LVIII (1998): 38-44
[31] Paul, "Keeping Track of the Time: Emending the Counterfactual Analysis of Causation," *Analysis*, LVIII (1998): 191-98
[32] "Trumping Preemption," this JOURNAL, XCVII, 4 (April 2000): 165-81.

is not, because the laws say that the first spells are the consequential
ones. Nevertheless, there is no counterfactual dependence of the
prince's becoming a frog on Merlin's spell, because Morgana's spell is a
dependency-breaking backup (*ibid.*, p. 165).

The potential ancestor theory is no help because "there is [no]
failure of intermediary events along the Morgana process (we may
dramatize this by stipulating that spells work directly, without any
intermediaries)" (*ibid.*). As for the hastener theory, there is no
"would-be difference in time or manner of the effect absent Merlin's
spell" (*ibid.*).[33] It seems that "nothing remains by which extant [coun-
terfactual accounts of causation] might distinguish Merlin's spell
from Morgana's in causal status" (*op. cit.*).

De facto dependence is established by holding fixed that *only
Merlin casts a first spell*, meaning by that that nobody else does. Had
Merlin held back, then given that other spells, if any, wait on his, no
spell would have been cast, hence the prince would have stayed one.
Of course, the possibility remains that a no less natural H can be
found that exposes the need as artificial. But as was argued in section
v, such an H is at odds with the intuitive judgment that Merlin's spell
meets the same need as Morgana's would have, if any actual event
does. And an H flouting that judgment would have, one suspects, to
be too unnatural to matter. But we need not settle this here; if (DF)
is not at an absolute loss when it comes to trumping preemption, then
it is at less of a loss than extant alternatives.

IX. CAUSATION AND INFLUENCE

Most extant alternatives, anyway. A theory recently proposed by
Lewis[34] can deal with all of the cases so far mentioned. Very roughly,
Lewis holds that c causes e if and only if c has variants c', c'',...and e has
variants e', e'',...such that which e-variant occurs depends on which
c-variant occurs. (That is direct causation; causation is the ancestral.)
Since the prince's fate is sensitive to the content of Merlin's spell but
not Morgana's, it is Merlin's spell that does the causing. But consider
a Merlin-type scenario due to Collins: "There is only one spell
[Merlin] may cast (standard prince-to-frog) and only one time of day
he may cast it (noon). No such limitations apply to Morgana.... The
prince's transfiguration is now in no way dependent on...how Merlin
acts.... Yet...it is still Merlin's spell, and not Morgana's, that causes the

---

[33] For related difficulties, see Hugh Rice, "David Lewis's Awkward Cases of
Redundant Causation," *Analysis*, LIX (1999): 157-64
[34] "Causation as Influence," this JOURNAL, XCVII, 4 (April 2000): 182-97.

prince to turn into a frog."[35] This appears to show that a cause need not influence its effect, while a noncause can influence the effect greatly.

Our task is to find a *G* modulo which the transfiguration depends on Merlin's spell rather than Morgana's, despite that the influence is all on the side of the latter. The same *G* used above appears to work here as well. Had Merlin held back in *G*-ish conditions, there would have been no spells at all, and so no transfigurations. If there is a more natural *H* portraying the need as artificial, it is not easy to find.

How much influence must a preemptive cause exercise over how its effect occurs? None, if the how is a how of *manner*. But suppose that it is a how of *means*. One feels that an event *greatly* redirecting the process leading to *e* is a cause, while one *slightly* redirecting it is not.[36] Consider

> "Flip": A track splits briefly into subtracks A and B and then reconverges. Further on lies the station. Which subtrack a train takes to the recon- vergence point is controlled by the position of a small electrical switch. Railroad worker Billy, bored with his job, has taken to flipping the switch back and forth when he hears a train approaching. Today's train would have used subtrack A had Billy's (final) act of switch-flipping not redi- rected it to B.

No doubt that final flip does cause the train to arrive by one route rather than the other. But does it contribute to the train's arriving as such? I take it the answer is that it does not. This might seem to create a problem for the DF theory, since the arrival depends on the final flip modulo a quite natural *G*. Holding fixed that nothing goes down subtrack A, there would have been no way forward had the switch not been flipped. Are we then not overgenerating, as Lewis undergener- ates? His account demands a kind of influence that is not essential to causation; DF fails to demand a kind of influence that is essential.

But this is to forget an important part of the theory. It is true that *G* puts the effect in need of that final flip. But the need is easily exposed as artificial, for example, by the null fact. The effect's fallback needs—its needs if the final flip had not occurred—are, let us say, for 100 one-yard motions down A. Does Billy's flipping of the switch meet any of these needs? Clearly not. Does it cancel any? No; they persist into the actual scenario as needs for 100 one-yard mo- tions down B. By the lights of the null fact, then, no real need is met by Billy's action but only an artificial one. (It might be thought that

---

[35] "Preemptive Preemption," p. 231.
[36] This is a particular theme of Hall, "Causation and the Price of Transitivity."

the null fact is no more natural than *G*. But if it makes a need artificial, then so do all "insensitive" facts: all facts *H* such that $\neg Ox >$ *H* whenever $\neg Ox > \neg Oe$.[37] *G* is sure to be less natural than some such fact.)

Note that, if we change the example so that fallback needs are cancelled, our causal intuitions change too. Suppose that subtrack A is damaged as the train approaches. The repair crew is working madly to fix it; they cry out to Billy "Flip your switch!" and he does. Now Billy's action plays a definite role. It does not matter if subtrack A is by some miracle repaired just in time. By relieving the effect of its need for that miracle, Billy's action makes itself eligible for the role of cause.

This is a good place to acknowledge that, although technically one fallback need is as good as another, in practice not all such needs are taken equally seriously. What if subtrack A is *normally* disconnected? The drawbridge it runs over is kept open to let boats through, except when sensors pick up an approaching train, at which times it routinely and automatically closes. The need that Billy's action cancels is, let us say, the need for a generally reliable mechanism to work the way it is supposed to. Such a need may be considered too slight to protect the action that cancels it from charges of artificiality. I am not sure what makes a need "serious" enough to escape this sort of criticism. One key factor seems to be this: how far-fetched the notion is of its not being met.[38] Far-fetchedness is a largely objective matter, to do with the prior probability of the need's not being met, and the remoteness of the scenarios where that happens. But nonobjective factors may play a role, too, such as how worried we are about the effect's not occurring for that particular reason, and how impressed we are by the size of the canceled need.[39]

---

[37] It is enough to show that NAN $\subseteq$ HAN, that is, every event that the null fact puts *e* in need of, *H* puts *e* in need of as well. Suppose that *x* belongs to NAN. Then $\neg Ox > \neg Oe$, which since *H* is insensitive means that $\neg Ox > H$. The last two counterfactuals imply that $(\neg Ox \,\&\, H) > \neg Oe$, whence *x* belongs to HAN. (As usual, *x* ranges over events occurring after the point at which the nearest *c*-free world branches off from actuality.)

[38] I take the notion of far-fetchedness from Collins, "Preemptive Preemption."

[39] This bears on an example of McDermott's: "Suppose that I reach out and catch a passing cricket ball. The next thing along in the ball's direction of motion was a solid brick wall. Beyond that was a window. Did my action prevent the ball from hitting the window? (Did it cause the ball to *not* hit the window?)" — "Redundant Causation," p. 525. Reactions to this change as interest rises in the need that gets canceled: the need would have been met by the ball's hitting the wall. (I assume that the need *is* canceled; for if it recurs it is met by the catch, whence the catch is a cause. It may help to replace the wall with a powerful steady wind.) Usually, we think, 'The wall was all set to stop that ball; it could

I spoke of the feeling that a c making "big" changes in the process leading to e is better qualified for the role of cause than a c making "small" changes in that process. If small changes are changes leaving the fallback needs more or less in place, the de facto theory bears this feeling out. The following is an example of Ned Hall's (somewhat abridged).

> "Kiss": Billy and Suzy meet for coffee. Suzy kisses Billy passionately, confessing that she is in love with him. Billy is thrilled—for he has long been secretly in love with Suzy, as well. Much later, as he is giddily walking home, he whistles a certain tune. What would have happened had she not kissed him? Well, they would have had their usual pleasant coffee together, and afterward he would have taken care of various errands, and it just so happens that in one of the stores he would have visited, he would have heard that very tune, and it would have stuck in his head, and consequently he would have whistled it on his way home...even though there is the failure of counterfactual dependence typical of switching cases (if Suzy had not kissed Billy, he still would have whistled), there is, of course, no question whatsoever that as things stand, the kiss is among the causes of the whistling.[40]

That seems right: the kiss is a cause of the whistling. But the example is not really typical of switching cases, for the kiss makes major changes in the process by which Billy comes to be whistling. The whistling's needs in the fallback (kissless) scenario are weighted toward the period *after* Billy leaves the coffee shop. They include Billy's deciding to drop into that particular store, the store's staying open until he arrives, the radio's playing that particular tune, and so on. All of these needs, and they are not trivial, are cancelled when Suzy gives Billy the kiss.

Suppose I am right that it is because Suzy's kiss takes the whistling out of a certain kind of late-afternoon jeopardy that we are willing to credit it as a cause. Someone might say: the kiss also puts the whistling into *new* jeopardy; it now becomes important, for instance, that Billy steers clear of a certain malicious gossip. Why is everyday causal

---

and would have done so easily; the catch gets no credit whatever for relieving the effect of so piddling a need'. Occasionally, we think 'The ball had a great deal of momentum; the job of redirecting that momentum is not made any smaller by the fact that the wall was more than up to it; kudos to the catch for making the job unnecessary'.

[40] "Causation and the Price of Transitivity."

thought not more bothered by that? If the role of a cause is to minimize needs, should not introducing new needs be just as bad as preserving old ones?

The answer to this is that the role of a cause is *not* to minimize needs; it is to meet needs. It is just that the need has to be of the right type: real rather than artificial. Taking the effect out of jeopardy is what c does, not to fulfill its mission as a cause, but to establish the bona fides of a particular need. That having been done, it is meeting the need that makes c a cause.

<div align="right">STEPHEN YABLO</div>

Massachusetts Institute of Technology