CRÍTICA, *Revista Hispanoamericana de Filosofía*
Vol. XXIX, No. 86 (agosto 1997): 107–119

COUNTERFACTUAL DEPENDENCE AND BROKEN BAROMETERS: A RESPONSE TO FLICHMAN'S ARGUMENT*

HELEN BEEBEE
Philosophy Program, Research School of Social Sciences
Australian National University

I

In a recent paper,[1] Horacio Abeledo presents a counter-example to David Lewis's counterfactual analysis of causation. The counter-example was first put forward by Eduardo Flichman[2] and was subsequently endorsed by Dorothy Edgington.[3] Both Flichman and Edgington use the counter-example to support the view that causation is unanalysable; Abeledo's response is rather different. He argues that Lewis's analysis can be saved, but only at a considerable cost: partly because the escape route requires but does not have independent motivation, and partly because

* I am very grateful to an anonymous referee for extensive and incisive comments on an earlier version of this paper, which is much improved as a result.

[1] H. Abeledo, "Lewis's Causation: An Almost Fatal Example", *Crítica*, XXVII, no. 81, 1995, pp. 79–100.

[2] E.H. Flichman, "The Causalist Program, Rational or Irrational Persistence?", *Crítica*, XXI, no. 62, 1989, pp. 29–53.

[3] D. Edgington, "Explanation, Causation and Laws", *Crítica*, XXII, no. 66, 1990, pp. 55–73.

it is not obvious that it can be applied to all cases that are structurally similar to Flichman's counter-example.

After briefly rehearsing Flichman's argument, I shall show that the putative counter-example which he presents is not in fact a counter-example after all. Thus there is no need for the costly alterations which Abeledo proposes on Lewis's behalf; nor is there any motivation here for Flichman's and Edgington's view that the concept of causation cannot be analysed.

The importance of showing why Flichman's argument does not succeed lies in the fact that if it were successful, it would scupper not only Lewis's own theory of causation, which in fact fails for other reasons,[4] but also more recent attempts to analyse causation using the counterfactual machinery which Lewis devised.[5] Flichman's objection is aimed not at the specific details of Lewis's theory of causation but at the notion of counterfactual dependence which is central to all its successors.

II

Lewis defines causation in terms of counterfactual dependence.[6] Roughly (and assuming determinism), event *c* causes event *e* if and only if *e* causally depends on *c* or there is a chain of causal dependence, via intermediate events,

[4] See P. Menzies, "Probabilistic Causation and Causal Processes: A Critique of Lewis", *Philosophy of Science*, 56, 1989, pp. 642–663.

[5] See for instance Menzies, 1989; H. Beebee, "Taking Hindrance Seriously", *Philosophical Studies*, 88, 1997, pp. 81–101; and M. Ramachandran, "A Counterfactual Analysis of Causation", *Mind*, 106, 1997, pp. 263–277.

[6] See D.K. Lewis, "Causation", *Journal of Philosophy*, 70, 1973, pp. 556–567, reprinted in his *Philosophical Papers*, vol. II (Oxford University Press, New York, 1986, pp. 159–172), for his original analysis of causation. I assume determinism for the purposes of this paper; see the Postscript to "Causation" in Lewis (1986), pp. 172–213, for his proposal for extending the analysis to indeterministic cases.

running from c to e. An event b causally depends on event a if and only if a and b are actual, distinct events[7] and the counterfactual "if a had not occurred, b would not have occurred" is true, i.e. if and only if

$$\sim O(a) \mathrel{\Box\!\!\rightarrow} \sim O(b)$$

is true. And Lewis's analysis of counterfactuals tells us that this counterfactual is true if and only if b fails to occur at the closest possible world(s) at which a fails to occur.

Since causation is asymmetric, Lewis needs the relevant counterfactuals to be asymmetric too: in general if $\sim O(a) \mathrel{\Box\!\!\rightarrow} \sim O(b)$ is true, $\sim O(b) \mathrel{\Box\!\!\rightarrow} \sim O(a)$ needs to be false if his analysis is to succeed.

Flichman's counter-example works as follows. We have the following events, all of which actually occur:

p: the atmospheric pressure being 1000mb
b: the barometer's working well
r: the barometer's reading 1000mb

p caused r, and $\sim O(p) \mathrel{\Box\!\!\rightarrow} \sim O(r)$ is true as Lewis's analysis requires. But r did not also cause p, so for Lewis's analysis to succeed,

(1) $\sim O(r) \mathrel{\Box\!\!\rightarrow} \sim O(p)$

must be false. And, according to Lewis, it *is* false: "If the reading had been higher, would the pressure have been

[7] "Distinct" here is a technical notion which does not simply mean "different". "We may take it", Lewis says, "that when one event implies another, then they are not distinct and their counterfactual dependence is not causal" (p. 256 of "Events" in his (1986), pp. 241–269). For example, my being tired and hungry is a different event to my being tired (since my forgetting lunch is a cause of the former but not the latter event), but the two events are not distinct, since my being tired and hungry implies that I am tired. So the counterfactual dependence of my being tired and hungry on my being tired is not causal dependence.

higher? Or would the barometer have been malfunctioning? The second sounds better: a higher reading would have been an incorrect reading."[8]

As Flichman points out, Lewis here seems to be claiming that (1) is false on the grounds that

(2) $\sim O(r) \; \Box\!\!\rightarrow \; \sim O(b)$

is true: if the barometer had had a different reading, it would not have been working properly. But, Flichman argues, if (2) is true then it follows from Lewis's definition of causal dependence that *r* caused *b*; and this, of course, is false.

Hence Lewis seems to face a severe problem: rejection of (1) seems to involve a commitment to (2), so he cannot consistently reject both (1) and (2). But this is just what his analysis of causation requires, since *r* caused neither *p* nor *b*.

III

My argument against Flichman falls into two parts. First (in section IV) I want to show that Lewis in fact has independent motivation for denying (1): independent, that is, of the truth or falsity of (2). So Lewis is not —despite what he seems to suggest the passage just quoted— required to assert (2) in order to motivate the denial of (1).

Of course, this leaves the status of (2) undecided. Lewis still has to face up to the objection that, even if the truth of (2) is not *required* in order for (l) to be false, the passage quoted above shows that he nevertheless holds it true; and this commits him to the false consequence that *r* caused *b*. The second part of my argument, in section V, shows that despite his assertion that if the reading had been different the barometer would have been malfunctioning, Lewis is

[8] Lewis (1973), p. 169.

not thereby commited to the truth of (2), for there is no way of specifying event b which makes (2) an acceptable formalisation of that assertion. So he can after all consistently reject both (1) and (2).

IV

First, then, I need to show that Lewis's theory of counterfactuals renders (1) false —and that it does so without making any presuppositions about the truth or falsity of (2). I shall give only a brief account; the finer details are to be found in Lewis (1979).[9]

To see whether or not (1) is true, we need to see which $\sim r$-worlds are closest to our own world. Closeness of worlds goes by similarity, and, roughly, similarity of worlds is to be judged according to how big the spatio-temporal region of perfect match of particular matters of fact is, and according to how well the worlds match with respect to their laws of nature.

Assuming determinism, any $\sim r$-world will *either* have been different from our world with respect to particular matters of fact since time immemorial, *or* have laws which, from our perspective, involve a "miracle": its laws, unlike ours, allow for r to fail to occur given at least some period of history that is identical to that of our own world. Lewis's criteria for evaluating similarity entail that worlds which preserve perfect match of particular facts for some period of history but involve a small miracle to stop r occurring are closer to ours than ones which involve no miracles but which never match our world with respect to particular matters of fact. And of the "miracle" worlds, those which preserve the spatio-temporal region of perfect match of

[9] See Lewis, "Counterfactual Dependence and Time's Arrow", *Noûs*, 13, 1979, pp. 445–476.

particular matters of fact for longest are closest to our own world.

Given this much, it's easy to see why the analysis renders (1) false. Some reasonably close $\sim r$-worlds will involve a miracle that makes the atmospheric pressure differ from 1000mb and thereby (since all other laws of nature remain the same as in the actual world) makes the barometer read something different to 1000mb. Others keep the atmospheric pressure the same but involve a miracle that directly makes the barometer give a different reading. The latter, p-worlds are closer to our own than are the former, $\sim p$-worlds, since the effect of atmospheric pressure on barometer readings is not instantaneous: it takes time (though perhaps hardly any) for the barometer to register the correct atmospheric pressure. So particular matters of fact at the former worlds depart from our own before the time of r, since the required miracle must happen a little before r. But the latter worlds can retain perfect match right up until the time of r, and hence are closer. And since the laws governing the weather are the same at those worlds as they are at ours, the atmospheric pressure will be the same at those worlds as it is at our own —1000mb. Hence (1) is false. Instead,

(3) $\sim \mathrm{O}(r) \mathrel{\Box\!\!\rightarrow} \mathrm{O}(p)$

is true: had the reading been different, the atmospheric pressure would have been just the same as it actually was.

V

What about (2)? Well, the important thing to notice is that Flichman's argument simply *assumes* that (2) is an adequate formalisation of —or at least follows from— Lewis's comment quoted earlier, which can be paraphrased as:

(4) If the reading had been different, the barometer would not have been working properly.

I take it, as Lewis does in the quoted passage, that (4) is true. If the reading had been different, the atmospheric pressure would still have been 1000mb. *A fortiori* the barometer would have failed to give the right reading, and would therefore not have been working properly. The question is whether, as Flichman assumes, commitment to the truth of (4) commits Lewis to the truth of (2). I shall argue that it does not.

Flichman's crucial assumption is that there is some event *b* which is 'the event corresponding to the fact that the barometer is functioning well'.[10] But the existence of such an event is not something which can just be assumed, since many facts —like the fact that the departmental seminar always happens on a Thursday, or the fact that pigs cannot fly— do not have events which "correspond" to them. Counterfactuals which involve the non-obtaining of facts like these cannot be recast as statements which assert that one *event* counterfactually depends on another; that is, they do not entail any statement with the logical form of (2). For example 'If pigs could fly, then the departmental seminars would not always take place on Thursdays' cannot be recast in the form '$\sim O(x) \;\Box\!\!\rightarrow \sim O(y)$', since there are no actual events x and y which "correspond" to the relevant facts.

So the question which needs to be asked is: is there any event *b* which corresponds to the fact that the barometer is working well, and which renders (2) true? My answer is going to be 'no': there is no legitimate way of characterising *b* which renders (2) true. Hence there is no way of making the crucial step from (4) to (2), and Flichman's argument against Lewis therefore fails.

[10] Flichman (1989), p. 34.

By characterising *b* as the event of the barometer's working well, Flichman characterises *b* as the event of the barometer's having a dispositional property: the property of being disposed to yield a reading of such and such mb when the atmospheric pressure is such and such mb. Are there, in general, events which consist in an object's having a dispositional property? According to Lewis, the answer is 'yes'; but it is a qualified 'yes', and it is the qualification that is crucial here. The qualification is that while events can be *accidentally* characterised in terms of dispositional properties, they cannot be *essentially* so characterised. (An essential characterisation of an event *e* is a characterisation which *e* cannot fail to satisfy: at every world where *e* occurs, it satisfies the characterisation. And an accidental characterisation of *e* is one which *e can* fail to satisfy: there are worlds where *e* occurs but does not satisfy the characterisation.)

Before returning to the issue of the correct essential characterisation of *b*, it is worth saying something about why Lewis thinks that dispositional properties ought to be banned from counting as essential properties of events. According to Lewis, the reason is that dispositional properties are both disjunctive and too extrinsic; and both disjunctive and overly extrinsic properties are in general banned by his theory of events from counting as essential characterisations of events. So the ban on dispositional properties falls out as a special case of a more general ban.

Why are dispositional properties disjunctive and overly extrinsic? For Lewis, dispositional properties are disjunctive because to say that an object has a dispositional property is to say that it has *some categorical basis or other* in virtue of which it is disposed to do such and such (e.g. give a reading of 1000mb) in such and such circumstances (e.g. the atmospheric pressure's being 1000mb). Since different objects can have the same disposition while varying with respect to the categorical basis of the disposition (there's more than one way to make a barometer), the disposition is the disjunction of all of its possible categorical bases. And dispositional properties are extrinsic to some extent, since whether or not an object has a particular disposition depends not just on its intrinsic nature —the categorical basis— but also on what the laws of nature are. Intrinsically identical objects can have different dispositions at different worlds.[11]

Why are disjunctive and overly extrinsic properties banned? Consider disjunctive properties first: "Fred talks, and his talking causes Ted to laugh. Suppose that besides Fred's talking there is another event, the disjunctive event of Fred's talking-or-walking. Without it, Fred's talking would not have occurred, and neither would Ted's laughing. So this disjunctive event also causes Ted to laugh. That is intuitively wrong. No such event causes Ted's laughing, or anything else. Given the theses I took as my starting point, that can only be because there is no such event. Hence disjunctive events are to be rejected."[12]

Overly extrinsic properties too give rise to counterfactual dependence where there is no causation; to use another example of Lewis's, without the widowing of Xanthippe the subsequent cooling of Socrates's body would not have occurred, since had Xanthippe not been widowed, Socrates would not have died.[13] Lewis's solution, again, is to ban descriptions like 'Xanthippe's becoming a widow' from counting as essential characterisations of events, so that the above counterfactual, though true, does not assert the counterfactual dependence of the cooling of Socrates's

[11] See Lewis, "Causal Explanation" in his (1986), pp. 214–240; Section III.

[12] D.K. Lewis, "Events" in his (1986), p. 267.

[13] See Lewis, "Events", p. 263.

body on any *event*; hence Lewis's theory of causation does not yield the erroneous result that Xanthippe's becoming a widow *caused* the cooling of Socrates's body.

What all this shows is that Lewis has a theory of events and a theory of dispositions which, when taken together, imply that b —which by hypothesis is an event which occurs at the actual world— cannot *essentially* be the barometer's being disposed to give the right reading, since on Lewis's account no such event occurs at this or any other world. Hence if Flichman's required entailment of (2) by (4) is to go through, b can only be *accidentally* characterised as the event of the barometer's working well.[14]

What, then, can event b essentially be? The only serious alternative is to construe b essentially as the barometer's having the internal structure it in fact does have —this being (at our world) the categorical basis for its disposition to give the right reading.

However, this construal of b renders (2) false, and for familiar reasons. Some close $\sim r$-worlds will be ones where a miracle happens so that the barometer's internal constitution changes, thus making it give a reading different to 1000mb. Others will not involve any change in the barometer's internal constitution at all: there will just be a miracle that stops r itself occurring. (Of course, in such worlds the barometer will not be working properly, since it gives the wrong reading. But on the current proposal the barometer's being disposed to give the right reading is only an accidental feature of b: b can still occur at worlds where, because the laws of nature are different, the barometer lacks the

[14] In fact there are reasons for denying that events can be essentially characterised in terms of dispositions that are independent of Lewis's ban on disjunctive and overly extrinsic events; see F. Jackson, "Essentialism, Mental Properties and Causation", *Proceedings of the Aristotelian Society*, XCV, 1995, pp. 253–268 (especially pp. 255–258).

disposition.) As with (1), one of the latter, *b*-worlds, will be closer than any of the former, $\sim b$-worlds, since the latter worlds can match the actual world with respect to particular matters of fact right up until the time of *r*. Hence (2) is false: if the barometer had given a different reading, its internal constitution would have remained the same, and hence, on the current construal of *b*, *b* would still have occurred.

It follows that if (2) is to follow from (4), *b* cannot be construed as the barometer's having the internal structure it has either, since on this construal (2) is false.

VI

I suppose there are many ways of giving an essential characterisation of "the" actual event which corresponds to the fact that the barometer is working well (which is to say, in Lewis's terminology, that there are many events satisfying this description, differing in respect of which of their properties are essential and which accidental).[15] But I see no reason to suppose that any of these events will be such as to make (2) follow from (4). In other words, I see no reason to think that there is any way of recasting (4) as a claim about the counterfactual dependence of some event *b* on *r*. Since it is (2) rather than (4) that generates Flichman's objection, the objection does not succeed: Lewis can consistently reject both (1) and (2). No suspiciously *ad hoc* rescue attempt, along the lines suggested by Abeledo, is needed; nor is there any cause here for alarm for those of us who still hold out for a successful counterfactual analysis of causation.

Recibido: 12 de abril de 1997

[15] See Lewis, "Events", *op. cit.*

RESUMEN

El artículo consiste en una defensa del análisis contrafáctico de la causación de David Lewis en contra de un argumento presentado por primera vez por Eduardo Flichman. El argumento de Flichman involucra una situación en la cual tienen lugar los tres sucesos siguientes:

p: una presión atmosférica de 1000mb
b: el funcionamiento correcto del barómetro
r: una lectura en el barómetro de 1000mb

Si el análisis de Lewis ha de tener éxito, la fórmula contrafáctica

(1) $\sim O(r) \mathrel{\Box\!\!\rightarrow} \sim O(p)$

debe ser falsa. Pero Lewis mismo justifica la afirmación de que (1) es falsa mediante la siguiente observación: "Si la lectura hubiese sido más alta, ¿habría habido una mayor presión? O ¿habría estado funcionando mal el barómetro? Lo segundo suena mejor: una lectura más alta habría sido una lectura incorrecta."

Flichman infiere de esta aserción que Lewis está comprometido con:

(2) $\sim O(r) \mathrel{\Box\!\!\rightarrow} \sim O(b)$

Se sigue del análisis de Lewis de la causación que r causó b; y esto, desde luego, es falso.

Sostengo que Flichman se equivoca al inferir que Lewis está comprometido con (2) a partir del hecho de que haga la afirmación citada más arriba. Flichman supone que hay cierto suceso b que corresponde al hecho de que el barómetro funcione bien. Pero de hecho la única manera de lograr que (2) sea verdadera es suponer que b puede caracterizarse esencialmente como el hecho de que el barómetro tenga una propiedad disposicional (a saber, la disposición de ofrecer la lectura correcta), y Lewis niega de manera explícita que las propiedades disposicionales puedan ser propiedades esenciales de los sucesos.

Por otra parte, si suponemos que la propiedad disposicional es simplemente una característica accidental de b, entonces (2)

es falsa. Por lo tanto, la afirmación de Lewis citada más arriba no puede interpretarse razonablemente como la afirmación de que cierto suceso b depende contrafácticamente de r; y, por lo tanto, la objeción de Flichman no funciona.

[*Traducción: Héctor Islas*]