Knowledge as Caused Belief

Wylie Breckenridge

I want to propose and begin to defend the following analysis of knowledge:

(CB) S knows that p iff S was caused by p to believe that p.

("CB" for "caused belief".) I do not know if such a theory has already been proposed. I have not yet read one in the literature, but there is much literature that I have not yet read. It first occurred to me while reading Nozick's counterfactual analysis of knowledge[1], which seems to have much in common with Lewis's counterfactual analyses of event causation[2]. It was reinforced in my mind while reading Dretske's suggestion that the notion of 'relevant alternatives' plays a role in our concept of knowledge[3], because I think that an analogous notion plays a role in our concept of an event, and therefore in our concept of event causation. One of my aims is to show that what Nozick and Dretske say about knowledge are just the things we might expect someone to say if (CB) is correct.

I

Lewis's analyses are of *event* causation – what it is for one *event* to be caused by another *event*. I want to take what we have learned from debates about those analyses and apply them to mine, so I want (CB) to be understood as a statement about event causation too. *Prima facie* it is not – it is about a state of affairs being the cause of a belief, and neither a state of affairs nor a belief is an event. So I really want the analysis to say this:

(CB') S knows that p iff the event of p coming to be the case caused the event of S coming to believe that p.

The picture I have in mind in this. The objects of our knowledge are states of affairs - to know that p is to know that the state of affairs p is the case. If p is the case then because, I will assume, there was a time at which p was not the case, there must have been an event of p coming to be the case. And if S believes that p then because, I will assume, there was a time at which S did not believe that p, there must have been an event of S coming to believe that p. So if S knows that p when we have these two events available to refer to. This picture is no doubt fraught with difficulties, but I will leave their meeting to be done elsewhere. On the understanding that (CB) is to be read as (CB'), I will keep it as the expression of my analysis.

II

(Comments about some of (CB)'s immediate consequences, such as S knowing that p implying both that p is true and that S believes that p. Comments about how (CB) handles the Gettier examples. Comments about how (CB) handles knowledge of truths that do not appear to have to causal efficacy, like mathematical and logical truths and truths about (merely) possible worlds.)

---

[1] See Nozick (1981).
[2] See Lewis (2000) for a summary of his various analyses.
[3] See Dretske (1970).

III

Nozick's counterfactual analysis of knowledge says that the following conditions are individually necessary and jointly sufficient for S to know that p:

(1) p is true,
(2) S believes that p,
(3) If p were not the case then S would not believe that p, and
(4) If p were the case then S would believe that p.

According to (CB), if S knows that p then S was caused by p to believe that p. So p must be the case and S must believe that p. Furthermore, the event of p coming to be the case must have caused the event of S coming to believe that p. According to the simplest of Lewis's analyses, where C and E are actual events:

(L) C caused E iff if C had not occurred then E would not have occurred.

So according to (CB) and (L) together, if the event of p coming to be the case had not occurred then the event of S coming to believe that p would not have occurred. That is, if p were not the case then S would not believe that p. Furthermore, Lewis assumes a theory of counterfactuals according to which if a and c are both true then it is trivially true that if a had been true then c would have been true. So if p is the case and S believes that p then if p were the case then S would believe that p. So according to (CB) and (L) together, (1) to (4) are individually necessary for S to know that p. Conversely, if (1), (2) and (3) are true then according to (L) p caused S to believe that p, so according to (CB) S knows that p. That is, according to (CB) and (L) together, (1) to (4) are jointly sufficient for S to know that p (in fact, (1) to (3) are jointly sufficient). So Nozick's analysis is a consequence of (CB) and (L).

I don't want to suggest that this is evidence in favour of the combined theory (CB)+(L), because I think that Nozick's analysis is wrong and only wrong analyses can have wrong consequences. I think that Nozick's analysis is wrong because there are circumstances in which S knows that p, but if p had not been the case then S would still have believed that p (insert example here). But I do want to suggest that the fault with (CB)+(L) lies not in (CB) but entirely in (L). In fact it is not hard to show that (L) is wrong (see Lewis (2000), pp. ?-?). If I am right in claiming that knowledge is to be analysed in terms of causation, then because there are some quite attractive (but false) counterfactual analyses of causation, we should not be surprised that there are some quite attractive (but false) counterfactual analyses of knowledge, like Nozick's. So my claim here is that the existence of analyses like Lewis's and Nozick's is evidence for the truth of (CB). I also want to claim that if (CB) is correct then epistemologists who hope for a counterfactual analysis of knowledge can sit back and let much of the work be done by metaphysicians who hope for a counterfactual analysis of event causation.

IV

Fred Dretske claims that to know that p is to know that p within a framework of "relevant alternatives" - it is to know that p rather than p_1, or p_2, or ... for some

"contrasts" p_1, p_2, etc.[4] We cannot change this set of contrasts without changing what it is that is said to be known, and by changing the set of contrasts we can change what is said to be known without changing the sentence that we use to express it. By shifting emphasis on the words in "S knows that Lefty killed Otto", for example, we can invoke different contrast sets and thereby alter what S is claimed to know. By saying "S knows that *Lefty* killed Otto" we can claim that S knows that Lefty (rather than George, Mike or someone else) killed Otto. By saying "S knows that Lefty *killed* Otto" we can claim that S knows that Lefty killed (rather than merely threatened or injured) Otto.

I think that a correct theory of events will say something along the same lines: that for event E to occur is for event E to occur within a framework of relative alternatives - it is for E to occur rather than E_1, or E_2, or ... for some contrast events E_1, E_2, etc. We cannot change this set of contrasts without changing what it is that is said to occur, and by changing the set of contrasts we can change what is said to occur without changing the sentence that we use to express it. We can use the sentence "Lefty killed Otto" to claim the occurrence of the event in which Lefty (rather than George) killed Otto, and also to claim the occurrence of the event in which Lefty killed (rather than merely injured) Otto. (Incidentally, one reason for me thinking that this is what a theory of events should say is that it would simplify, I believe, counterfactual theories of event causation like Lewis's. I hope to argue this elsewhere.)

If I am right in claiming that events occur always within a framework of relevant alternatives, then it follows from (CB) that to know that p is always to know that p within in a framework of relevant alternatives. That is, Dretske's claims are a consequence of (CB). Why? If S knows that p then according to (CB) S's belief that p was caused by the event of p coming to be the case. But according to the relevant alternatives view of events, the event of p coming to be the case is the event of p (rather the p_1, or, p_2, or ...) coming to be the case, so S's belief that p was caused by the event of p coming to be the case, rather than p_1 coming to be the case, or p_2 coming to be the case, etc. So S knows that p rather than p_1, or p_2, etc. Because Dretske's claims are a consequence of (CB) and an (independently motivated) relevant alternatives view of events, I take it that any strength of Dretske's theory is also a strength of (CB).

V

Dretske's theory is one of several "contextualist" theories of knowledge, each of which claims that what it takes to know that p depends upon the context in which a knowledge claim is made. Another such theory claims that knowledge comes in degrees, and that what it takes to know that p depends upon the context because the context determines the standard for knowledge - the degree to which S must know that p in order for "S knows that p" to be true. This gives rise to one response to the sceptic. Following DeRose (1995), if we let 'H' stand for some sceptical hypothesis and 'O' stand for some proposition we would ordinarily claim to know, then we can take the sceptic's arguments to have this form:

1. I don't know that not-H
2. If I don't know that not-H then I don't know that O

So,

C. I don't know that O

[4] See Dretske (1970), esp. pp. 143-144.

The perplexing thing about sceptical arguments (for well chosen H and O) is that they seem to be valid and the premises seem to be true but the conclusion seems to be false. Contextualists who claim that knowledge comes in degrees can argue that we are wrong to take premise 1 as being true and at the same time to take the conclusion C to be false, because to do so requires us to evaluate each according to a different standard for knowledge. If we fix the standard for knowledge, then if 1 is true then C is also true. If I think it's true that I don't know that I'm a brain in a vat, then I must be evaluating my knowledge according to a standard high enough to make it true that I don't know that I have hands. Contrapositively, if I think it's true that I know that I have hands, then I must be evaluating my knowledge claim according to a standard low enough to make it false that I don't know that I'm a brain in a vat.

Does (CB) allow room for such a contextualist response? I could argue that causation comes in degrees. I could argue, for instance, that both (i) the cow on the track and (ii) the interesting show on TV last night that kept the farmer up late and lead him to sleep in and be late to round up the cow for milking, were both causes of the train's derailment, but that the cow was a cause to a greater degree than the TV show. Then I could argue that if (CB) is right and also causation comes in degrees, then knowledge too comes in degrees, thus allowing for a contextualist response like the one just considered.

But I'm inclined to think that causation does not comes in degrees, and that neither does knowledge. At least not the knowledge that (CB) purports to be an analysis of. I'm inclined to say that the knowledge claimed in premise 1 is a *different type* of knowledge to that denied in the conclusion C. Indeed, if causation is an all-or-nothing affair then I think I *have* to say something like this to account for the intuitive plausibility of 1. If (CB) is right, then I ought not to accept that 1 is true. I certainly believe that I'm not a brain in a vat, so to think that 1 is true is to think that this belief was not caused by the state of affairs of my not being a brain in a vat. But I don't think that. I agree that it's possible, but I don't think that it's actual. So if (CB) is right then I should grant the sceptic that 1 could well be true, not that it *is* true. Unless I'm thinking of knowledge in a different way when I think that 1 is true. And that's what I'm inclined to say. Specifically, I'm inclined to say that I'm thinking of knowledge as subjective certainty when I think that 1 is true, and as caused belief when I think that C is false. That is, I have two distinct concepts of knowledge, only one of which does (CB) purport to analyse. I do believe that I have hands and I do believe that this belief was caused by my having hands, so I do believe that in one sense (the "caused belief" sense) I know that I have hands. But I also believe that I might be wrong, and that I am not subjectively certain that I have hands, and hence that in another sense (the "subjective certainty" sense) I don't know that I have hands. The sceptical argument trades on an equivocation between these two senses of 'knows'.

(Insert here further arguments for why we have these two concepts of knowledge, so that my claim that we do does not seem ad hoc.)

VI

Finally, I want to discuss what (CB) says about the closure of knowledge under known logical implication. If S knows that p and S knows that p implies q then,

according to (CB), does S thereby know that q (assuming that S believes all the logical consequences of her beliefs)? This is a tricky question, and I'm not sure what answer to give. Rain falls on my roof and thereby causes me to believe that it is raining and thereby to believe that it is cloudy (because rain implies cloud). Do I know that it is cloudy? According to (CB), I know that it is cloudy iff my belief that it is cloudy was caused by its being cloudy. Was it? The relevant events are (C) it becoming cloudy, and (E) my coming to believe that it is cloudy, and the question is, did (C) cause (E)? It would be wrong to argue that (C) caused it to rain which caused me to believe that it is raining which caused (E), and thereby conclude by following the causal chain that (C) caused (E). This would be wrong simply because it would be wrong to claim that (C) caused it to rain. It being cloudy is a logical consequence of it raining, but it raining is not a causal consequence of it being cloudy. Nevertheless, I'm inclined to think that (C) did cause (E), largely on the basis of considering what would have been the case if (C) had not occurred (that is, if it had not become cloudy) - in that case, I think that (E) would not have occurred (that is, I would not have come to believe that it is cloudy). I am not assuming the truth of any counterfactual analysis of causation here, just noting that counterfactual considerations do play a role (and in this case a large role) in my causal judgments.

But I do I know that it is cloudy because it is a logical consequence of something else that I know (that it is raining), or because of some other feature of this example (such as some non-logical connection between cloud and rain)? Suppose I look at my hands and thereby come to believe that I have hands and thereby to believe that I am not a brain in a vat. Do I know that I am not a brain in a vat? Here the relevant events are (C) it becoming the case that I am not a brain in a vat, and (E) my coming to believe that I am not a brain in a vat. It is fairly easy to think about event (E) and what might have caused it (looking at my hands, for example, is intuitively a cause). But event (C) is hard to think about. When and how did it happen? If there is a causal chain of events from (C) to (E) then what might they be? Suppose that many years ago God chose to create us as we take ourselves to be, rather than as brains in vats, and that event (C) was this creation. Did (C) cause (E)? That is, did God's creation (way back then) cause me to believe (now) that I have hands? I'm inclined to think not. His creation *enabled* me to form this believe, but it did not *cause* me to form it. But then I don't know a logical consequence of something that I do know (that I have hands), which suggests that if (CB) is right then the closure of knowledge fails.

But this needs more consideration. (Include also a discussion of what to say if (CB) does deny closure - for example, of how closure can be false and yet seem so attractive.)

References

F. Dretske (1970), 'Epistemic Operators', in DeRose and Warfield, pp. 131-44.
K. DeRose (1995), 'Solving the Skeptical Problem', in DeRose and Warfield, pp. 183-219.
K. DeRose and T. A. Warfield (1999), eds., *Skepticism: A Contemporary Reader*.
D. K. Lewis (2000), 'Causation as Influence', *Journal of Philosophy* **97**, pp. 182-97.
R. Nozick (1981), 'Philosophical Explanations', in DeRose and Warfield, pp. 156-79.