

CAUSAL RELEVANCE¹

Stephen Yablo
Massachusetts Institute of Technology

An astonishing number of philosophers hold that mental states are poorly positioned to cause behavior. Timothy Williamson in recent work distinguishes two sources of doubt about the causal prospects of *wide* mental states in particular.²

One is that causal attributions are supposed to have “an appropriate generality” (81).³ The idea goes back at least to Hume, who asserts in Section XV of Book I of the *Treatise*, “Rules by which to judge of causes and effects,” that

where several different objects produce the same effect, it must be by means of some quality, which we discover to be common amongst them. For as like effects imply like causes, we must always ascribe the causation to the circumstances, wherein we discover the resemblance (174).

Enter now the Twin Earth examples. You do the same thing—drink—whether it is water (H₂O) you desire or twater (XYZ). This seems to be just what Hume was talking about. Different “objects” (water-desire and twater-desire) produce the same effect (drinking). The causation must be ascribed to some quality that is common amongst them. That common element would seem to be something intrinsic: a narrow-content state, or a piece of brain-writing, or perhaps even a brain state.

A second source of doubt about wide mental causes is this. Causes operate via causal mechanisms; so they should be where the mechanisms are, specifically, where they are set into motion. (This idea goes back to Hume too, in his emphasis on contiguity.) Since behavior issues from mechanisms internal to the agent (“the causing of my present action is here and now” (65)), its causes should be internal as well.

[N]arrow conditions must play a privileged role in the causal explanation of action. If a causal explanation of action cites a broad mental condition, an

underlying narrow condition must do the real work. We can isolate that narrow condition by subtracting from the broad mental condition the environmental accretions that make it broad... (65).

It seems to strengthen the case against wide causation that the same conclusion is reached from two almost opposite perspectives. One argument *zooms in* on the causal transaction, on the theory that whatever sets the machinery into motion can be seen in its entirety from close up. The other argument *pans out* so that the transaction appears against the background of other transactions of the same sort. That extrinsic factors can be varied indefinitely at no cost to the effect exposes them as irrelevant hangers-on.

All of this is to remind us that defending wide causation is a non-trivial task.⁴ Mistakes will have to be found in both of the above arguments: *generality* and *locality*, I will call them. And they will have to be distinct and independent mistakes, since the arguments take such very different approaches. Williamson appears to do this. His complaint about *generality* is that it rests on a false assumption.⁵

The narrow explanation is *not* always the more general. Take my seeing the ball on a certain occasion and the internal state of my head on that occasion. It may be that the internal state is as sufficient for the catch, other things equal. But it is nowhere near as necessary. I could have seen the ball in any number of ways, intrinsically speaking, without that compromising my ability to make the catch. I could have been watching it through either eye, for instance, or with my head cocked at any of a large number of angles. If the generality of an explanation is a matter of how well the factors cited correlate with the effect, then the narrow explanation is in this case the less general.

Against *locality* he says that

The most the argument shows is that other things equal, the real cause is not my knowing, or widely believing, but the internal core thereof: the “organismic contribution” as Dennett used to say. Other things are far from being equal, however, for states of knowing (or widely believing) do not *have* internal cores to speak of. Knowing *would* have an internal core if it were “composite” = the conjunction of an internal fact with an external one. But knowing is “prime.” If you insist on splitting knowledge states up between internal and external components, it is a huge disjunction of such conjunctions. The answer to “how do you have to be *internally*, to know there is water nearby?” is “no particular way; what you need is to be in the right kind of harmony with your environment.” If you tried to whittle knowledge states down to their environment-independent internal core, you would be whittling them down to nothing.

As promised, each argument runs into a distinctive sort of trouble. *Generality* fails because wide mental states oftentimes *correlate* better. *Locality* fails because wide mental states are typically *prime*.⁶ This tidy arrangement proves unstable, however, for the response to locality is incomplete.

If Williamson is right, then a certain popular *strategy* for finding intrinsic surrogate states does not work; one cannot carve off the extrinsic parts and expect to have anything useful left. But why should the *real* causes of behavior have to be reachable by performing some kind of bracketing operation on the *alleged* causes? Maybe the alleged causes are wrong through and through. To get to the real causes you have to drop them entirely and look somewhere else. It is clear there are *some* intrinsic surrogates, for there are brain states. Why not assign causal responsibility to them?

Now, Williamson does have something to say against the causal ambitions of brain states:

If one cites a sufficient condition for the condition to be explained...the purported explanation can nevertheless fail because the condition to be explained would still have obtained in the same way even if the cited condition had not obtained.... Many features of [the brain state] will be quite irrelevant to the obtaining of [the effect]. They will concern physical events that form no part of the causal chain between the agent's initial mental state and the final performance of the action. The agent would have performed the action anyway, even if those features had been different (81–2).⁷

Brain states are faulted here on the score of generality. They may be sufficient, but they are nowhere near as necessary, for their alleged behavioral effects.⁸ Switching to (partly extrinsic) mental causes buys us a lot of necessity at the cost of not much sufficiency.

Isn't this a good and convincing reply? It is a good reply to something. But remember, the worry was about locality: no action at a distance. How are *correlational* considerations supposed to affect the *metaphysical* thought that extrinsic states bring in factors too far away to make a causal difference?

Of course, Williamson does not suppose for a moment that every correlation is causal. He says, "The high correlations between prime mental conditions and conditions on subsequent action constitute *defeasible* evidence for the causal effectiveness of the prime conditions" (88, emphasis added). But the route from correlation to causation is represented as pretty short: "Higher correlations constituting a genuinely rival explanation would be needed to defeat that evidence" (88).

The problem is that some high correlations are non-causal for reasons other than the one Williamson mentions, viz. that a higher correlation exists constituting a genuinely rival explanation.⁹ One can, after all, get a perfect correlation by taking the disjunction of all conditions sufficient for the effect. It is hard to do better than perfect, yet the correlation isn't causal. So there have got to be other defeaters than trumping by rival correlations "constituting a genuinely rival explanation."

Someone worried about the locality problem will say: "Here is my opening. I suggest that your knowledge correlation is non-causal for one of these other reasons. It is non-causal because it is too extrinsic; it incorporates elements too far away to make a causal difference."

This is not such an unmotivated thing to say. A little generality is causally speaking a good thing: being hit by a bus, Williamson notes, is a better candidate for cause of death than being hit by a *red* bus. But there are limits. The pursuit of greater and greater generality eventually takes one *away* from the cause. Williamson acknowledges this in places:

Some purported explanations achieve spurious generality by using disjunctive concepts. For example, if someone was crying because she was bereaved, it does not improve the explanation to say that she was crying because she was bereaved or chopping onions. But ordinary mental concepts of prime conditions (such as the concept of seeing) are not disjunctive (83).

I agree that ordinary mental concepts are not disjunctive. But how do we know they do not possess some *other* feature that defeats their claim to feature in genuine explanations?

Not all "spurious generality" has its source in disjunctiveness. If you are looking for a property of liquids that correlates well with the property of unplugging clogged drains, it would be hard to improve on *plumber-recommendedness*. Plumbers keep track of the substances currently getting stuck in drains, and the types of solvent available, and they recommend accordingly. Plumber-recommendedness certainly scores higher than any *chemical* property that might be mentioned; for there are various chemical preparations that work about equally well.

And yet plumber-recommendedness does no causal work. The various chemical properties do it, notwithstanding their poorer correlation with the effect. The objection to plumber-recommendedness is not that it is disjunctive but that it brings in factors too far away to influence goings-on in the drain. That could be how it is with knowledge and the brain state; knowledge correlates better, the brain state does the work.

So far, then, we lack an answer to the metaphysical charge that wide mental states bring in factors too far away to bear on the causation happening here and now. I see only three lines of response with any chance of success. The first is *denial*: "Who says the factors making a state wide are too far away?" The second is *dismissal*: "Let me tell you my theory of causal relevance; it says wide mental states can be causes." The third is *diagnosis*: "Here is why the locality argument *seems* right." The rest of the paper looks at examples of each strategy.

Denial

Williamson sometimes suggests that the factors making a mental state wide are *not* too far away, or at least won't *continue* to be as the time of action draws closer.

One is thirsty; how likely is one to be drinking soon? Likely enough, if one sees water. Much less likely, if what one sees is a mirage... Concepts of broad mental conditions give us a better understanding of connections between present states and actions in the non-immediate future, because the connections involve interactions with the environment (75).¹⁰

Having water in view lets me advance towards the water, renewing my perceptual link to it at various points along the way. It is not just that my perceptual state is at once wide and effective; it is effective *because* it is wide. Advancing on the water, I draw causally on the factors that make my state one of seeing water, as opposed to merely seeming to see it or seeing watery stuff. For I draw causally on the water itself, and on my continuing perceptual rapport with it.

I *said* that I draw causally on the factors that make my state wide. But, truth be told, it is really only some of them I draw on. That it is *water* I am seeing is owing in part to the stuff's being H₂O as opposed to XYZ. But as far as my behavior is concerned, it might as well have been XYZ. That it is *seeing* I am up to depends on the process by which my visual experience is renewed. That process must be not only reliable but "of the right type" (no deviant causal chains). As far as my behavior is concerned, though, deviant causal chains would be fine, as long as they made for a replenishable supply of good information on the same topic.

So it does not seem that *all* the factors making the broad state broad make themselves felt en route to the action.

A second example is this. Knowing that the mine contains gold, you dig until you find it. Your belief that it contained gold does not explain your persistence as well as your knowing does. For your knowing involves *inter alia* that your belief is based on true considerations and not false ones. (This is the no false lemmas condition on knowledge.) Suppose your evidence had been that there was gold behind a Maltese Cross carved into the cave wall. You would have looked there first, and abandoned the search on not finding anything. Your persistence was due in part to your not concentrating all your hopes on that particular spot. Here again, extrinsic aspects of your knowledge are not too remote to make a causal difference.

But again, that one such aspect plays a causal role doesn't mean that all do. You would (also) not have known that there was gold in the mine if some misleading testimony given in Carson City—testimony you should have been aware of but weren't—had not been refuted by court

records in Reno—with you again unaware of the fact. The court records in Reno play no role in your continued digging here. But they are a factor in your knowing. This seems to bear out the localist complaint that your knowing incorporates factors too far away to affect the course of events.

Dismissal

This is the strategy where we attempt an analysis of causal relevance, and show that broad mental conditions are relevant in the terms of that analysis. The locality argument is not refuted; it is not even mentioned; it is overruled by a higher court. We try to establish by other means the conclusion that locality was supposed to threaten.

I had Williamson claiming that states of mind are likelier causes of behavior than brain states because they are better correlated with behavior.¹¹ But his words can be taken another way. He starts out by saying that a “purported explanation can . . . fail because the condition to be explained would still have obtained in the same way even if the cited condition had not obtained” (81). (Psst . . . note the counterfactual.) This applies in particular to action explanations that pin the blame on lower-level features of the agent’s brain: “The agent would have performed the action anyway, even if those features had been different” (82). (There’s that counterfactual again.)

To judge by these passages, the brain state is rejected not for correlational reasons but counterfactual ones.¹² This could help, because counterfactuals, more anyway than correlation coefficients, appear to have a direct causal significance. Almost everyone’s first thought about causal relevance is

(CT)

A property P of *x* is causally relevant to effect *y* iff
y would not have ensued, had *x* occurred without P.

Call that the *counterfactual theory of causal relevance*. What does it say about the cases of concern to us here?

Had I not seen that water, I would not now be drinking. So, according to the counterfactual theory, my seeing the water was causally relevant. It cannot be said, in most cases, that I would not be drinking had I not earlier been in a rather precise neural condition. So my precise neural condition is not causally relevant.

There is only one problem with this, which is that the counterfactual theory is wrong. This is clear from the debate about the causal relevance of *moral* properties.

Gilbert Harman said roughly this: there is no knowledge of moral properties unless moral beliefs are caused by such properties. When you look at particular cases, though, it turns out to be the underlying natural properties that do the causing. There is no need to appeal to the cruelty of setting cats on fire to explain our belief that those (cat-burning) children are being cruel. We came *into* the situation believing that torturing cats was cruel. Our belief that the children are being cruel is sufficiently explained by the non-moral fact that (cat-torturing) is what they are doing.

Nicholas Sturgeon asked why it should make moral properties *irrelevant* that non-moral properties are *sufficient*. After all, the moral properties might supervene on, or be otherwise bound up with, the non-moral ones. Cat-burning, to stick with that example, is necessarily cruel, or close enough for present purposes. But then, had the children not been behaving cruelly, they would not have been burning a cat, whence (unless Plan B was to *pretend* to burn a cat) their behavior would not have induced the belief that the children were being cruel. There would have been no belief, then, had there not been cruelty. How then can the cruelty be considered irrelevant to the belief?

Because Sturgeon argues from counterfactual dependence to causal relevance, responses to Sturgeon are a good source of counterexamples to (CT). Consider the case of Donald (due to Judy Thomson). To relieve his boredom at a particularly dreadful talk,

[Donald] suddenly shouted Boo! at the speaker. In consequence, there was a loud Boo! sound on the tape recording of the speech. Now if there are moral facts at all, there is such a moral fact as that Donald's shouting Boo! was rude. But that fact was surely epiphenomenal relative to the . . . presence of a Boo! sound on the tape: the fact of Donald's shouting's having been rude surely plays no role at all in explaining the fact of the Boo! sound . . . [And yet] the case passes the counterfactual test for operativeness . . . For we may suppose that if Donald's shouting . . . hadn't been rude, it wouldn't have been [a case of] his shouting Boo at the speaker in mid-speech . . . in that case there would have been no shout at all during the speech . . . and . . . therefore no Boo sound on the tape.¹³

That is pretty convincing, I think. But it doesn't offer any guidance about how to fix the theory; and so we turn to a couple of structural problems. The first is that an effect that depends counterfactually on P depends all too often on $P \vee Q$ as well, even if Q is quite irrelevant. The second is similar, except that now it is $P \wedge Q$ that inherits "relevance" from P.

Parasitic Disjunction

Had Donald's Boo! not been loud, there would have been no Boo sound on the tape. So there is no Boo sound on the tape in the nearest

world w where it fails to be loud. Assuming that the Boo! would not have been regretted the next day had it not been so loud, w is also the nearest world where the Boo! fails to be *loud or regretted the next day*; whence there would (also) not have been a Boo sound on the tape had the Boo! lacked the disjunctive property of being loud or regretted the next day. Its property of being loud or regretted the next day therefore helped it (says the counterfactual theory) to register on the tape.

Parasitic Conjunction

There is no Boo sound on the tape in the nearest world w where Donald's Boo! fails to be loud. Assuming that Donald would have delivered this quieter Boo! from the same position at the back of the room, w is also the nearest world where the Boo! fails to be both loud and a long way from the microphone. Since this is by hypothesis a world where there is no Boo sound on the tape, the tape recorder would not have picked it up had the Boo! lacked the conjunctive property of being *loud and far from the microphone*. Its property of being loud and far from the microphone therefore helped it (says the counterfactual theory) to register on the tape.

What is going on in these cases? One reason *loud or regretted* seems irrelevant is that there is a stronger property *loud* on which the effect still depends. But that cannot be all that is wrong, for the Boo! was also over ten decibels. It is true of *over ten decibels* too that there is a stronger property *loud* on which the effect still depends. And that the Boo! was over ten decibels seems highly relevant to the tape recorder's picking it up.

Is the problem that *loud or regretted* is not as natural as *loud*? No, because that doesn't bother us when the weaker property is obtained by chipping away aspects of the stronger one that the effect would not have missed. It does not seem at all irrelevant to its registering on the tape that the Boo! was *loud enough given its distance from the microphone*.

The problem with *loud or regretted* is that it exhibits both of these failings at once. Passing from *loud* to *loud or regretted*, we suffer a decline in naturalness with no compensating gain on the score of attunement with the effect. (The effect would not have occurred had the Boo! been loud or regretted but not loud.) *Loud or regretted* is not merely weak but *egregiously, pointlessly*, weak.

The conjunctive property of being *loud and far from the microphone* offends in the opposite way. Passing from *loud* to it, we pile on irrelevancies—loud is better proportioned to the effect than the property that replaces it—with no compensating gain on the score of naturalness.¹⁴ This is a property that is not merely strong but *egregiously* strong.

Now let's try to make these notions a bit precise. Suppose that we are trying to identify the causally relevant properties of some object or event x ,

and that Q_- and Q_+ are weaker and stronger properties of x ; Q_+ necessitates Q_- but not conversely. Egregiousness is defined in two steps:

Def Q_- is $\left\{ \begin{smallmatrix} \text{better} \\ \text{worse} \end{smallmatrix} \right\}$ proportioned to y than Q_+ iff y would $\left\{ \begin{smallmatrix} \text{not} \\ \text{still} \end{smallmatrix} \right\}$ have occurred, had x possessed Q_- but not Q_+ .

Def A property P of x is egregiously $\left\{ \begin{smallmatrix} \text{weak} \\ \text{strong} \end{smallmatrix} \right\}$ (relative to effect y) iff some $\left\{ \begin{smallmatrix} \text{more natural stronger} \\ \text{as natural weaker} \end{smallmatrix} \right\}$ property of x is better proportioned to y than P is.¹⁵

Consider now the *proportionality theory of causal relevance*:

(PT)

A property P of x is causally relevant to effect y iff

- (a) had x lacked P , y would not have occurred¹⁶
- (b) P is not egregiously weak or strong.

Loud or regretted meets condition (a) but not (b). That *loud* is a more natural stronger property better proportioned to the effect means that *loud or regretted* is egregiously weak. Similarly *loud and far from the microphone* meets (a) but not (b). It is egregiously strong, since *loud* is a no less natural weaker property better proportioned to the effect.

How does my seeing the water fare on this theory? The worry would be that it is egregiously strong. That is, my seeing the water is screened off¹⁷ by a weaker property that is no less natural. The obvious choice is my seeming to see the water and its really being there. (I assume this is no less natural than my seeing the water, or at least that the objector could consistently maintain that it is.) Would the drinking still have occurred, had I only seemed to see water which was in fact there?

It is hard to feel sure about this; one wants to know *why* the seeming to see wasn't real seeing. Had I veridically hallucinated the whole time, my progress towards the water would not have been much affected. Had I veridically hallucinated for a brief moment between periods of non-veridical hallucination, it would have been another story. It does seem clear, though, that the drinking *might well* not have occurred, had I only seemed to see water which was however there. The seeing is not egregiously strong unless I *would* have had the drink had I merely seemed to see what was really there. That I might not have had it shows that the would-claim is false. So the seeing is not egregiously strong.

Or consider my knowing that you are home; the effect is my knocking and knocking for ten minutes. The worry would again be that it is egregiously strong, that is, screened off by a weaker property that is no less natural. The obvious choice is my truly and rationally believing that

you are home. Would I have knocked for ten minutes, had I correctly and with reason believed that you were home, without knowing you were home?

Once again, it depends on why the knowledge was missing. But it does seem that I *might* not have knocked so long; for I might have failed to know because of overlooking the “Not Home” sign posted on your door. Unless I somehow continued to overlook it for ten whole minutes, the effect (knocking for ten minutes) would not have occurred. And that is all Tim needs. The knowing is not egregiously strong unless I *would* have kept on knocking for ten minutes had I failed to know while still rationally and truly believing. That I might have stopped knocking shows that it is not the case that I would have kept on. So the knowing is not egregiously strong.

Now the rude Donald example. Thomson asks what properties of the *shouting* are relevant, but that raises a question we would rather avoid, namely, would the effect still have occurred had the shouting lacked a property that is essential to it (loudness, say)? Better to ask about properties of the *shouter*. Donald has at a particular time the property of behaving rudely. How relevant is his possession of that property to the sound on the tape?

Unlike my knowing that you were at home, which we worried might be egregiously strong, Donald’s rudeness is under suspicion of being egregiously weak, that is, of being weaker than some more natural property that it fails to screen off. A stronger property would be an action-type φ such that necessarily, all φ ing is rude. Suppose for argument’s sake that φ ing is shouting Boo!, or shouting Boo! in a crowded room. Boo!-shouters are all doing roughly the same thing, while (as any Grade 3 teacher will tell you) there is no end to the forms that rudeness can take. Shouting Boo!, even shouting it in a crowded room, is a whole lot more natural than the property of putting on some sort of rude performance of other.

One further thing is required for Donald’s rudeness to be irrelevant because egregiously weak. Donald’s property of shouting Boo! in a crowded room must be better proportioned to the effect—to the Boo sound on the tape—than his property of behaving rudely is. Would there still have been a Boo sound on the tape, had Donald been rude otherwise than by shouting Boo! in a crowded room? Well, consider some nearby alternatives. Would there still have been a Boo sound on the tape had Donald in a crowded room rudely shouted something *other* than Boo!/? Or had he rudely *whispered* Boo! in a crowded room? Or had he rudely shouted Boo! in a room containing just him and the speaker? On the face of it, there would *not* have been a Boo sound in these scenarios. So on the face of it, the rudeness is egregiously weak =_{df} up against a stronger, more natural, property that is better proportioned to the effect.¹⁸

Diagnosis

A reason has been given for rejecting the locality argument: its conclusion is at odds with a plausible-seeming analysis of causal relevance. But we have yet to find an actual mistake in the argument. Its central claim, that there are extrinsic factors in knowing too far away to affect the outcome, seems true enough. A factor in your knowing there is gold in the mine is that some misleading testimony given on the subject in Carson City was refuted by court records in Reno, all without your knowledge. The court records in Reno play no role in your continued digging here. But they are a factor in your knowing. And so there are factors in your knowing that make no difference to the event your knowing is said to cause. Doesn't this show that your knowing was more than the effect needed and so not its cause?

That depends. "More than the effect needed" might mean that the knowing has elements without which the effect would still have occurred. In that sense, the knowing is more than the effect needed. In that sense, though, *every* cause is more than the effect needs. You can always find irrelevant aspects to take out. It is just that after a while, the gains in relevance are more than offset by the loss of naturalness. Concern about this puts a natural brake on the process of whittling away irrelevancies: a process that, left unchecked, would lead to every effect being blamed on the fact that things occurred sufficient for an effect just like that one.

To say that a proposed cause involved "more than the effect needed" might, on the other hand, mean that it involves irrelevancies that can be whittled away *with no loss of naturalness*: it might in other words mean that the proposed cause is *egregiously strong*. If that is what is meant, though, then it just does not follow, from the fact that your digging was insensitive to goings-on in Carson City and Reno, that your state of knowing the mine to contain gold was more than the effect needed. It might be that, although there are Carson-city-indifferent states that screen the knowing off from the effect, none of them is as natural than the knowing. It might be that the strongest *natural* weakening of the knowing is too weak to do the job itself.

Now we can see where the locality argument goes wrong. The problem does not lie with the assumption that behavior has intrinsic causes. Maybe it does and maybe it doesn't; the defender of wide causation need not take a stand on this. Nor should we be bothered by the internalist's claim that wide "causes" are liable to contain extrinsic irrelevancies. They *are* liable to contain extrinsic irrelevancies. The mistake is to run these thoughts too closely together, maintaining that they contain extrinsic irrelevancies *because* they make irrelevant additions to internalist-style intrinsic causes. Wide causes contain irrelevancies because all causes do, including narrow causes if such there be. The internalist's larger mistake is to forget that proportionality is not pursued at all costs but traded off against naturalness.

It seems hardly open to doubt that wide mental conditions effect *one* attractive such tradeoff—one local maximum of the relevant utility function. Internalists are welcome to search for a second local maximum more to their liking.

Notes

1. This paper was written for a conference on Tim Williamson's *Knowledge and its Limits* held at University College, London, and presented subsequently at the 2001 Sofia conference in Veracruz, with comments by Louise Antony. Thanks to Louise, Alex Byrne, Tim Crane, Paul Horwich, Elizabeth Fricker, Michael Martin, David Papineau, Ralph Wedgwood, and Tim Williamson for extremely helpful comments and criticism.
2. Williamson 1998 and 2000. The focus here will be on propositional attitude states, counting perception as a propositional attitude. These can be wide either because of the content involved (*realizing that water is refreshing*) or the attitude taken toward that content (*realizing that water is refreshing*). It is width due to the attitude taken that primarily interests Williamson, but for our purposes the cases can be lumped together.
3. Williamson references are to the book unless otherwise indicated.
4. Other recent defenders include Gibbons 2001, Jackson and Pettit 1988, and Yablo 1997.
5. What follows is my account of what I take to be Tim's position. Similarly for the response to locality.
6. Primeness enters in a limited way into the first response too, in that a prime state is touted as better correlated with the effect.
7. I have substituted brain states for Tim's actual target in this passage because the same considerations seem to apply.
8. Williamson favors a probabilistic reading of sufficiency and necessity.
9. Williamson certainly knows this. The perfect correlation example about to be given is his, and he comments about it that "we are willing to sacrifice some degree of correlation" for the sake of unified (non-disjunctive) correlata (89).
10. "In deliberating, one assesses alternative courses of action in light of one's beliefs and desires, decides which is best, and forms the intention to pursue it. . . . How and whether one puts the intention into effect depend on one's interaction with the environment in the intervening period" (1998, 396).
11. I said that this was not decisive, for too much correlation can be a *bad* thing causally speaking.
12. I doubt that this is Williamson's idea, not least because he told me it wasn't. Still, it would serve his interests if it worked.
13. Harman and Thomson 1996: 81
14. If a gain in naturalness is required to compensate for excessive strength, why not also to compensate for excessive weakness? (All we ask of the worse-proportioned weaker property is that it not be *less* natural than the better-proportioned stronger one.) There is a reason for this. An excessively strong property has aspects on which the effect does not depend, while an excessively weak one merely fails to include material on which the effect does depend. Including material that wouldn't have been missed is much more destructive of overall causal relevance than omitting

material that would have been missed. (With causal sufficiency it is the other way around; omitting material that would have been missed is “worse” than including material in whose absence the effect would still have occurred.)

15. This definition puts naturalness ahead of proportionality in the following sense: no amount of the latter can compensate for even a small loss of the former. A different definition would allow proportionality to be traded off—at a large discount, I assume—against naturalness. (This note was prompted by an example of Alex Byrne’s.)
16. Counterfactual theories of causal notions are subject to a standard objection, namely, the counterfactual fails if another cause would have taken over in the actual cause’s absence. I take preemption worries seriously, but this is not the place to discuss them. See Yablo 2002.
17. One property of x screens off another iff x could have possessed the first property without the second, and had it done so, the effect would still have occurred.
18. There are other examples where the case for egregious weakness/strength is harder to make out. Sometimes a property escapes being egregiously weak only by incorporating “elsewhere” material irrelevant to the effect, and/or escapes being egregiously strong only omitting “elsewhere” material relevant to the effect. So, it might be that *loud-or-regretted and far from the mike*, although not egregiously weak or strong, would be egregiously weak if not for the *far*, and egregiously strong if not for the *regretted*. This suggests a refinement of the (PT) account. P has *egregiously weak aspects* iff some better proportioned and no less natural P– is egregiously weak. P has *egregiously strong aspects* iff some better proportioned and no less natural P+ is egregiously strong. A property is causally relevant iff (a) the effect depends on it, (b) it is not egregiously weak or strong, and (c) it has no egregiously weak or strong aspects, *Loud-or-regretted and far from the mike* violates (c) twice over. It has egregiously strong aspects, since *loud-or-regretted* is (i) better proportioned to the effect, (ii) no less natural, and (iii) egregiously weak. It has egregiously strong aspects, since *loud and far from the mike* is (i) better proportioned to the effect, (ii) no less natural, and (iii) egregiously strong. This might be the right account for *behaving rudely* as well. It has egregiously strong aspects because for some suitable action-type ϕ , *rudely ϕ ing* is (i) better proportioned to the effect, (ii) no less natural, and (iii) egregiously strong since screened off by just ϕ ing.

References

- Braun, D. 1995. “Causally Relevant Properties,” *Philosophical Perspectives* 9, 447–475
- Jackson, F. and P. Pettit 1988. “Functionalism and Broad Content,” *Mind* XCVII, 381–400
- Gibbons, J. 2001. “Knowledge in Action,” *Philosophy and Phenomenological Research* 62, 579–600
- Harman, G., and J. J. Thomson. 1996. *Moral Realism and Moral Objectivity* (Oxford: Blackwell)
- Williamson, T. 1998. “The Broadness of the Mental: Some Logical Considerations,” *Philosophical Perspectives* 12
- Williamson, T. 2000. *Knowledge and Its Limits* (New York: Oxford University Press)
- Yablo, S. 1992. “Cause and Essence,” *Synthese* 93, 403–449
- Yablo, S. 1997. “Wide Causation,” *Philosophical Perspectives* 11, 251–281
- Yablo, S. 2002. “De Facto Dependence,” *Journal of Philosophy* 99, 130–148