

Ramachandran's Four Counterexamples

PAUL NOORDHOF

In my original paper, the diagram on page 99 had an inhibitory axon between *b* and *c* that should not have been present.

Murali Ramachandran has kindly provided me with four (alleged) counterexamples to the theory of causation which I recently put forward in *Mind* (Ramachandran 2000; Noordhof 1999). Space is limited for a response. Since this note will be published Ramachandran's paper, I will not set out the cases he gives. I refer the reader to the appropriate descriptions. I will also presume knowledge of the framework of my paper and just give page references in case this is helpful. I will try to couch the discussion with as little reference to the technical apparatus of my paper as possible. Maybe the ideas will stand out more clearly that way.

1. Cases 3 and 4

The crude idea behind my account is that causes raise the chances of their effects. The theory is just an attempt to refine this idea by reflecting upon various cases. One verdict, with which I agreed, was that mere delayers of a particular event were not causes of it (see Noordhof 1999, pp. 108–15, for discussion of delayers and hasteners). Obviously, in putting matters this way, I am allowing that a particular event may occur at other times than the time it actually occurred.¹ The difficulty is that a mere delayer very much raises the chance of an event occurring at the time it did. So instead of just focusing on the time the event actually occurred, I suggested that we compare the chance of the event occurring at the time it did, given the delayer had occurred, against each chance that the event had of occurring at each of the times given the delayer

¹ On page 109, I explain what would happen if I abandoned this assumption (Noordhof 1999).

had not occurred. A mere delayer fails significantly to raise the chance of the event occurring over the maximal background chance it had of occurring anyway (see Noordhof 1999, pp. 109–13 for further discussion and pp. 100, 104, 197 for more details on how the chances should be assessed).

In the case of the bomb, the problem for me is supposed to be this (Ramachandran 2000, pp. 311–2). Flicking switch *B* will not come out as a cause of the bomb's explosion. If switch *A* is flicked on first, the chance of the bomb exploding one second later is 0.5. This is greater than the chance of the bomb exploding after *B* is flicked on, when the bomb is in state 2, the chance being 0.2. The apparatus I used to get the right result for mere delayers allegedly gets the wrong result in the case of the bomb.

I considered a similar case in my paper: compound *X* with a critical time period (Noordhof 1999, pp. 114–5). I agreed that my proposal could not get the kind of conclusion that Ramachandran thinks is intuitive, but I explained how I could get the equivalent of the following: flicking switch *B* is a cause of the bomb's explosion *while it was in state 2* (hereafter, I will translate the treatment of the case I did discuss to Ramachandran's case). I think this minor modification of our intuitions by my proposal is justified by the fact that it represents the best non-arbitrary resolution of the difference between causing an event to occur and causing it to occur at a time. In allowing for the possibility of this departure from our intuitions, I am assuming a worst-case scenario for myself in at least two respects (Noordhof 1999, p. 114). First, I assume that the same particular event could occur as an explosion of the bomb in state 1 and as an explosion of the bomb in state 2. I would have no problem if the explosion of the bomb while in state 2 were a different token event from the bomb's explosion in state 1. Second, I assume that there are no events in the bomb which I could put in while the bomb is in state 1 which are significantly responsible for the high chance of the explosion then (I did not make this assumption explicit before) (for discussion of the role of the set, see Noordhof 1999, pp. 100–8). If one could, without distorting the transition from state 1 to state 2, there would be no problem. Since Ramachandran claims that the case he has described constitutes a counterexample to my approach he needs to make it plausible that these explosions are one and the same particular event and that there were no other events than flicking switch *A* which I might have put in . This he does not do. He also fails to explain how the case he provides introduces importantly new features that render the original discussion irrelevant. Since the burden of my argument was that the minor revision of our intuitions

was theoretically motivated, I think he needed to address that point (Noordhof 1999, pp. 114–5).²

In the case of the window opening, I do not share Ramachandran's view (Ramachandran 2000, pp. 312–3). He thinks it is clear that the inadvertent opening of the window is not a cause of the window shattering. He appears to reason as follows “the brick would have hit, and shattered, the window even if *c* [opening the window slightly] had not occurred—the impact would simply have occurred a bit later” (Ramachandran 2000, p. 312). The truth of the claim that the window would have shattered anyway depends upon the surrounding circumstances. The shattering would not have occurred without the inadvertent opening of the window if the brick were otherwise going to fall just off-target or be deflected at the last minute. My account tried to articulate the intuition that an event is a cause regardless of what *other* potential or actual causal chains were present. Our verdicts on early and late cases of preemption illustrates this intuition. My account also sat well with Penelope Mackie's characterisation of what made hasteners causes (in contrast with delayers). She suggested that hasteners are causes because they prevent events from occurring later *by bringing them about earlier*. By contrast, delayers bring events about later by preventing them from occurring earlier (Mackie 1992, pp. 493–5). Ramachandran's case is fairly described as the slight opening of the window preventing the shattering of the window later *by bringing it about earlier*. It also seems to fit the general pattern of the familiar case of early

² A natural thought to have is that one might limit the times at which one assesses the chance of explosion without the flicking of switch *B* to only those for which the flicking of switch *B* has some chance-raising or chance-lowering influence. The problem with this is that one will always have to add the time the alleged effect actually occurred to the set of times whether or not the flicking of switch *B* had an influence on the chance of the event at that time. If the explosion had a high chance of occurring then anyway, the flicking of switch *B* might come out as a cause because, although it had no influence then, at the times at which it did, the chance of the explosion occurring was very low. Unfortunately, if we do add the actual time to the set of times, when we shift to counterfactual situations as a result of putting certain events in \mathcal{C} , the *influence* of flicking switch *B* on the chance of the explosion at that time may well change due to the changed circumstances. This will lead to dubious verdicts on whether flicking switch *B* is a cause of the explosion. Even if this problem could be rectified, I suggest we shouldn't try. If there were some non-obvious limitation on the set of times which worked, I would argue that we should appeal to the notion defined in my proposal rather than continue working with a notion of causation for which it was hard to see the rationale (I remember Ramachandran putting a proposal to me along the lines discussed in this footnote back in early 1998).

preemption which, according to Ramachandran, is a clear case of causation (Ramachandran 1997, pp. 269–70).³

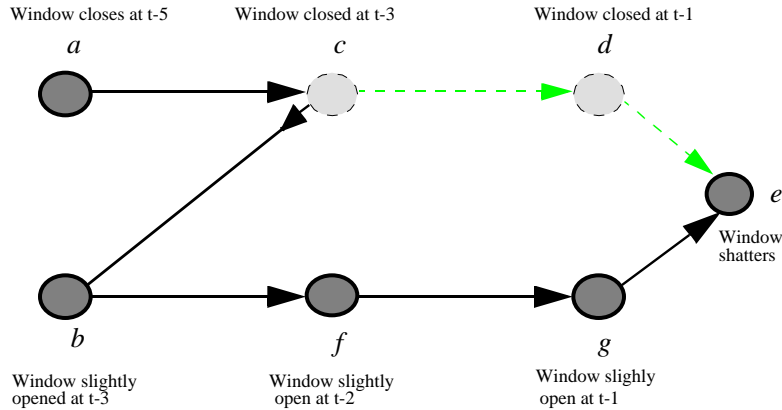


Figure 1⁴

The same reasoning that would declare the slight opening of a window not a cause—viz “the brick would have hit, and shattered, the window even if *c* [opening the window slightly] had not occurred—the impact would simply have occurred a bit later”—would make all cases of early preemption fail to involve a cause. For these reasons, I will stick with the verdict my proposal seems to deliver.

2. Cases 1 and 2

What I think are Ramachandran’s two more substantial and interesting objections raise questions about the similarity weighting of possible worlds that governs the assessment of the relevant counterfactuals. Ramachandran’s cases fail as counterexamples because he does not go into this matter and, once one does, it seems that the most plausible thing to say helps my theory.

³ With characteristic honesty, Ramachandran notes that earlier accounts he has defended (some along with me) also pronounce the slight window opening to be a cause. He does not say that he withdraws his judgement that, in a case of early preemption with the structure in Figure 1, *b* is a cause (Ramachandran 2000, p. 313).

⁴ There are many intermediate events between the nodes but the original neuron diagram did not rule that out either. One salient point of difference is that, in addition to the inhibitory connection between *b* and *c*, the events in the bottom row imply that the events in the top row are absent. However, I don’t see why this difference would justify a different verdict.

My theory rules out transitive indeterministic causation. I was convinced that causation was not transitive by Michael McDermott's paper, as indeed was Ramachandran (McDermott 1995; Ganeri, Noordhof and Ramachandran 1998). Ramachandran claims that, as a result, my theory is unable to deliver the verdict that the firing of *c* is a cause of the firing of *e* in the neuron chain case (Ramachandran 2000, pp. 310–1). If the firing of *c* had not occurred, there still is a slight chance that the firing of *d* (the intermediate neuron) would have occurred anyway. If the firing of *d* did occur, then the chance of *e* firing would be just as high as it would have been if the firing of *c* had occurred. So it would not be true that

If the firing of *c* were not to occur, then for any time *t*, it would be the case that, just before *t*, the chance of *e* firing would be at most *y* (where *y* is very much less than the chance of *e* firing if the firing of *c* had occurred). (Noordhof 1999, pp. 109–10)

He suggests the only solution is to embrace transitivity, which I don't want to do (Ramachandran 2000, p. 311).

I concede that, if *c* did not fire, there is still some chance that *d* fires. However, Ramachandran needs more than that to establish his point. He needs the following to be true.

- (I) If there is some chance that *d* fires even when *c* does not fire, then in one of the closest possible worlds in which *c* does not fire, *d* does fire.

That is because to assess what *at most* the chance of *e* firing would be, given *c* did not fire, we look in all the closest possible not-*c*-firing-worlds and take the greatest chance of *e* firing in such a world as the ceiling. It is the claim that *d* does fire in one of the closest not-*c*-firing-worlds that I reject.

Neither Ramachandran, nor I for that matter, discussed the similarity weighting of worlds that goes into the assessment of counterfactuals. I took myself to be appealing to Lewis's formulation.⁵

- (1) It is of the first importance to avoid big, widespread, diverse violations of law.
- (2) It is of the second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails.
- (3) It is of the third importance to avoid even small, localized, simple violations of law.
- (4) It is of little or no importance to secure approximate similarity of particular fact, even in matters which concern us greatly (Lewis 1979, pp. 47–8).

⁵ Ramachandran also seems to take himself to be working within Lewis's framework (e.g. see Ramachandran 1997, p. 263).

Lewis's approach faces a problem with indeterministic causation which the famous Nixon case brings out nicely (as he notes, see Lewis 1986, pp. 58–65).

If Nixon had pressed the button, there would have been a nuclear holocaust

appears true. However, in an indeterministic world, one could maximise the spatio-temporal region throughout which perfect match held by supposing that the signal from button to bomb, the visual effect of seeing Nixon pressing the button, Nixon's memory at having done so, and so on, all fail because they rely upon indeterministic processes. The world could converge back to our world directly after the button depression. The result would be that the following would be proclaimed true.

If Nixon had pressed the button, there would not have been a nuclear holocaust.

Even if the first counterfactual is not clearly true, the second is clearly false.

This suggests that we must add to Lewis's similarity weighting for possible worlds. The following seems plausible.

- (2) It is of second importance to minimise *differences* of particular fact from the actual world which are improbable.

It would not do to try to minimise all improbable particular facts because, as Lewis himself points out, some actual particular facts are improbable yet we would want to take them across to the closest possible worlds.⁶ We just don't want to *introduce* more improbabilities than necessary. It is highly improbable that the wire from button to bomb should fail to transmit a signal if the button is pressed. That's why we suppose that it does transmit the signal in the closest possible world and hence that it is false that, if Nixon had pressed the button, there would not have been a nuclear holocaust.

The same point applies to Ramachandran's sequence of neurones. The unadjusted similarity weighting would make it true that, if *c* hadn't fired, *d* would still have fired. This is clearly wrong. However, the proposed modification to the similarity weighting would get the intuitive verdict that this counterfactual is false. The fact that *d fires after c fails to fire* is highly improbable. So, to minimise the introduction of improbabilities, we suppose that *d* does not fire in the closest possible worlds.⁷ However,

⁶ For a similar reason, I don't think that Lewis's own solution works. He suggests that worlds with remarkable miracles—very unlikely events conspiring to form a pattern—are very unlikely are own (Lewis 1986, pp. 60–2). I don't think this is true. We hear talk of amazing coincidences all the time. I am prepared to allow this may be well grounded.

if that is the case, then my theory also yields the verdict that c 's firing is a cause of e 's firing after all.

It seems to me that, even if the proposal just made does not capture precisely the similarity weighting with which we work, something along these lines is needed. I think that it is more than likely that this will deny Ramachandran support for (I). Hence, I do not think he has established that there is a counterexample to my theory which I may only resolve by asserting transitivity.⁸

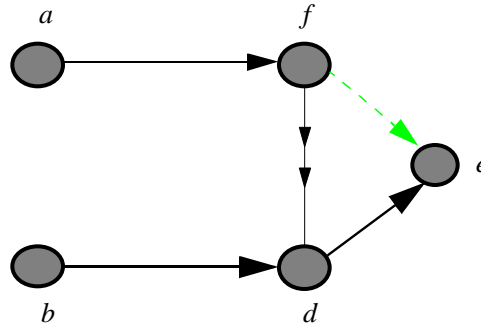
I will finally turn to Ramachandran's claim that my condition (IV) is misapplied. Since Ramachandran notes that the condition is not easy to grasp, I will begin by trying to make the thought behind it a little clearer.

The first three conditions of my theory establish whether an event, e_1 , significantly raises the chance of an event, e_2 , at the time e_2 occurred, over what I called the maximal background chance (Noordhof 1999, pp. 110, 120). Because the theory allows for the chances to be assessed in slightly different circumstances (for instance, when preempted chains are absent), the theory faces a problem with what I dubbed catalysts and anti-catalysts. These speed up or slow down a particular reaction. This allows for the following possibility. A chain of events may fail to complete in time to be a cause of e_2 because of the activity of an anti-catalyst. However, if the anti-catalyst were removed, then the chain of events would have completed, so causing e_2 at the time it actually occurred. The chain of events which actually failed to complete would satisfy the first three conditions of my account. Hence they are not sufficient for causation. In my original paper,

⁷ The firing of d itself is an actual event. So I am not claiming that it should be ruled out because it is improbable. As I wrote, what is improbable is the fact that d fired after c failed to fire.

⁸ My suggestion as to the similarity weighting of possible worlds needs a lot more discussion and defence. For some reasons why one must deny that if d 's firing has some chance of occurring then it is not the case that it would not occur, see Lewis (1986, pp. 64–5).

I provided the following illustration of such a case (Noordhof 1999, pp. 115–6).



If we consider circumstances in which the firing of d is absent, then the firing of a will satisfy conditions (I) to (III) with respect to the firing of e . Yet, the firing of a is not actually a cause of the firing of e . Condition (IV) provides a test of whether a causal chain has *actually* completed. Since the absence of d 's firing obviously distorts the situation, it, and events like it, must be present in the appropriate counterfactual circumstances. I tried to identify the feature of events like d 's firing that made their presence necessary.

My suggestion was that the problem stems from the fact that d 's firing plays two roles: it significantly raises the chance of e firing over the maximal background chance and it slows down the a – e process. Normally, we exclude an event which significantly raises the chance of an event, e , occurring, over the maximal background chance, to reveal how other events raise the chance as well. We assume that this will not affect the relationship between these other events and e firing. When an event combines two roles in the way that the firing of d does, this assumption is not legitimate. That is what the constraint on the membership of A brings out (Noordhof 1999, p. 117). It shows that the firing of d has an effect on the relationship between a 's firing and e 's firing at the same time it did even if we take away the fact that it independently raises the chance of e 's firing (that is, its satisfaction of conditions (I) to (III)).

Ramachandran's objection to my use of condition (IV) is that, if I am supposing that d 's firing may fail to satisfy conditions (I) to (III), I cannot assume that the other events in the set-up above continue to satisfy conditions (I) to (III). He claims that the other events will be involved in the *same kind of process*. The change in laws necessary to make d 's firing fail to satisfy conditions (I) to (III) will make the other events fail to satisfy them as well (Ramachandran 2000, p. 310).

I think his worry is overstated. First, it is not quite accurate to say that the events in question are part of the same kind of process. The firing of d is an anti-catalyst as well as a chance-raiser of e firing. Why should we assume that, in its chance-raising aspect, it works in the same way as other events which are not anti-catalysts? Second, Ramachandran assumes that the most similar set of laws to the laws which hold in the actual world would be one which abandons the law, or laws, actually governing that kind of process completely, rather than adjusting it, or them, so that just d -placed events cease to satisfy conditions (I) to (III) of my account. I can think of many of reasons why this assumption is mistaken. The most obvious one is that anti-catalytic events would have features which interact with their chance-raising features and the simplest set of laws would be ones which make an exception for them. Third, one needn't even rely upon a change of laws, for d 's firing to fail to satisfy conditions (I) to (III). The firing of d could occur in a sufficiently different time or place that it no longer stands in the required relationship to the *particular* event of e 's firing. Alternatively, there could be a novel event which inhibited d 's firing from raising the chance of e firing in some way. These options would certainly seem to be favoured by the similarity weighting of possible worlds set out by Lewis.

In spite of this, it now seems to me that I need not have formulated the test condition in the way that I did in terms of an event failing to satisfy *any of* conditions (I) to (III). Instead I might have required just that the event fail to satisfy *all of* conditions (I) to (III). The crucial difference is that, if I adopt the latter, then an event could fail to satisfy them merely by failing to raise the chance of e_2 at the time e_2 occurred. It is very unlikely that we would ever need to suppose a difference in the relevant process-governing laws to achieve that.

As a result, I do not think that Ramachandran has provided any counterexamples to my theory. That does not mean that I think my theory is unassailable. Like Ramachandran's own theory, it gives us a liberal notion of cause which will not be satisfactory to those who require something beefier (see Ramachandran 1998, p. 468). It awaits a proper resolution of the debate about the similarity weighting of possible worlds. Some theories of the individuation of events will be more congenial than others. Finally, if there were reason to suppose that causation is an inexplicable primitive, it would be possible to provide cases in which chance-raising, however refined, occurs yet it is stipulated that causation is absent. I remain to be convinced that my proposal faces straightforward counterexamples without detailed considerations drawn from one of these areas. So, although the issues Ramachandran raised were typi-

cally ingenious, sophisticated and challenging, I'm afraid I remain unmoved.⁹

Department of Philosophy
University of Nottingham
University Park
Nottingham NG7 2RD
UK
Paul.Noordhof@nottingham.ac.uk

PAUL NOORDHOF

REFERENCES

- Ganeri, Jonardon, Noordhof, Paul and Ramachandran, Murali 1998 "For a (revised) PCA Analysis". *Analysis*, 58, pp. 45–7.
- Lewis, David 1979: "Counterfactual Dependence and Time's Arrow", in his *Philosophical Papers*, Vol. 2, Oxford: Oxford University Press, 1986, pp. 32–66. Originally published in *Noûs*, 13, pp. 455–76.
- 1986: *Philosophical Papers*, Vol. 2. Oxford: Oxford University Press.
- Mackie, Penelope 1992: "Causing, Delaying, and Hastening: Do Rains Cause Fires?". *Mind*, 101, 403, pp. 483–500.
- McDermott, Michael 1995: "Redundant Causation". *The British Journal for the Philosophy of Science*, 46, pp. 323–44.
- Noordhof, Paul 1999: "Probabilistic Causation, Preemption and Counterfactuals". *Mind*, 108, pp. 95–125.
- Ramachandran, Murali 1997: "A Counterfactual Analysis of Causation". *Mind*, 106, pp. 263–77.
- 2000: "Noordhof on Probabilistic Causation". *Mind*, 109, pp. 309–13.

⁹ Although I am clearly an ungrateful recipient, I offer my thanks to Murali Ramachandran for making me think about issues that I would otherwise not have the wit to have thought about and, indeed, for taking the time to discuss my theory in the first place.