

Bilking the bilking argument

Rebecca Roache

[Forthcoming in *Analysis* 69.4 (October 2009)]

Is it conceptually possible for an event, *L*, to be the cause of an earlier event, *E*? Some writers have employed the so-called bilking argument to attempt to show that the idea of such backwards causation is incoherent (Flew 1956, 1957, 1964; Black 1956). According to this argument, if we are presented with what someone claims to be a case of backwards causation, it would be possible in principle to wait for *E* to occur, and then intervene to prevent the occurrence of *L*, thus demonstrating that *E* could not have been caused by *L* after all. Moreover, if our attempts to bilk *L*-type events (for short, *L*-events) having observed *E*-type events (*E*-events) always fail, we have grounds to argue that any causal relationship between the two is not one of backwards causation, but of ordinary, forwards, earlier-to-later causation.

Does the bilking argument succeed in showing backwards causation to be incoherent? Before answering this question, let us deal with a far simpler reason for deeming backwards causation incoherent, whose elucidation now will be useful later. To those attracted to the view that temporal order is determined by causal order¹, backwards causation may seem incoherent because on this view a cause is by definition earlier than its effect. I do not take issue with this view, and wish my conception of backwards causation to be compatible with it². By ‘backwards causation’, I mean causation that runs in the opposite direction to some other causal processes, such that the temporal order entailed by such backwards causation is the reverse of that entailed by those other causal processes. In most philosophical discussions of backwards causation, we are at least implicitly given reasons to believe that these other causal processes dictate the dominant temporal order in that world, with respect to which causation in the opposite direction occurs backwards. This is either because these other causal processes make up the vast majority of causal processes in that world, or because they include the mental causal processes of subjects with whom we are invited to identify, such that backwards causation runs

¹ See, for example, Hans Reichenbach (1958) and D.H. Mellor (1998).

² The view does not itself entail any particular conclusion about the possibility backwards causation. Whilst Mellor argues that backwards causation is impossible (1998: chapter 12), Reichenbach concedes that it could occur (1958: 139-43).

backwards with respect to the temporal order we experience. On the other hand, in worlds where roughly as many causal processes—including mental causal processes—run in one direction as in the other, it would be arbitrary to take either group of processes to determine the dominant temporal order. Since the group of worlds in which backwards causation is possible should include the latter sort of worlds, it is perhaps preferable to dispense with talk of backwards causation in favour of talk of bi-directional causation. But, since so many writers use the term ‘backwards causation’, I will follow suit.

Let us turn to the bilking argument. A full and satisfactory refutation of this argument would involve not only arguing that it does not show backwards causation to be impossible, but also addressing the attendant issues about agency; for if backwards causation is possible, then there is some sense in which we are not—even in principle—free to prevent *L* once *E* has occurred. Such curtailments of freedom would not show that backwards causation is impossible; they may show only that agency in worlds where backwards causation occurs differs from agency in worlds where it does not occur. Where one aims only to refute the claim that backwards causation is impossible, however, one can make a neater job of responding to the bilking argument if one is able to dispense with issues about agency. One writer has done this by considering only bilking scenarios that do not involve agents (Ismael 2003). Others—as we will see—have done so by constructing scenarios in which backwards causation occurs without the possibility of interference from agents bent on bilking (Dummett 1993³; Tooley 1997). Since backwards causation, if possible, is possible regardless of its implications for agency, I will respond to the bilking argument by imagining a scenario in which agents are free to attempt to bilk. I will not, however, draw out the implications for agency.

Let us begin by considering two similar attempts to respond to the bilking argument. Michael Dummett asks us to imagine a world very much like ours except that processes seem to occur in reverse, and in which we are present only as ‘intelligent trees’, able to observe but unable to intervene in what we see. In such a world, Dummett notes that ‘we should have great difficulty in arriving at causal explanations that accounted for events in terms of the processes which had *led up to*

³ Whilst Dummett (1993) describes a scenario in which agents are unable to interfere in backwards causal processes, he also goes on to consider in detail another scenario in which agents are able to attempt to bilk. Since one of his purposes in considering the latter scenario is to examine the implications for agency, I will not discuss it here.

them' (1993: 118). For example, in observing the reverse of the process in which apples fall from apple trees and roll across the ground before coming to a halt, we could not explain why the apples start rolling in a particular direction at a particular moment, and why they unfailingly then attach themselves to *apple* trees rather than to some other object. In order to explain such things, 'we should have to substitute a system of explanations of events in terms of the processes that led back to them from some subsequent moment' (118). That is, we should suppose that causation in such a world happens backwards with respect to the causal processes involved in our—the observers'—thought and experience. And, of course, since we observers are unable to intervene in the causal processes we observe, we are unable to demonstrate by bilking that these processes do not occur in reverse.

Michael Tooley imagines something similar:

Consider ... the following world. It consists of two spatial regions segregated from each other by a wall with some remarkable properties. First, the wall has always existed. Secondly, it is indestructible, and so it will exist at every future time. Thirdly, there is no way of getting around the wall, so that something can travel from the one region of space to the other only by going through the wall. Finally, the wall's properties differ depending upon the direction through the wall: in one direction, no causal processes at all can be transmitted from the one region to the other, while, in the opposite direction, light waves, but nothing else, can pass through the wall.

Suppose now that you are living on the side of the wall that can receive light waves from the other side, but from which no causal influence at all can be transmitted back in the opposite direction. On your side of the wall, life is very similar to the actual world. But on the other side, things are somewhat different. For example, next to you, on the other side, is a golf course. (Tooley 1997: 64)

What we observe happening on the golf course runs along similar lines to what we observe as Dummett's intelligent trees: golf balls spontaneously take off from the ground, fly through the air, collide with golfers' clubs, and settle neatly on the ground at their feet. Such observations, Tooley tells us, lead us to suppose that the causal processes on the other side of the wall run backwards relative to those on our side. And, since we cannot causally influence the events on the other side, we cannot bilk the apparent later causes of earlier events.

Do Dummett and Tooley describe cases of bilk-proof backwards causation? Hanoch Ben-Yami thinks not. He notes that, whilst Tooley's wall is intended to prevent the observers from interfering in the apparently backwards causal processes occurring on the other side, this constraint on their interference is 'contingent relative to the nature of the causal processes involved' (Ben-Yami 2007: 455). This means that, whilst Tooley has stipulated that the wall is indestructible, it is nevertheless the case that *if* the observers could break through it, they could bilk the alleged later causes of earlier events. Given this, it hardly matters whether or not the wall is indestructible, since

the nature of causality on the wall's observed side is the same, whether or not we interfere in the processes occurring there ... So it follows that there is no backwards causation ... in the processes on the observed side of the wall, whether or not there is any interference in these processes by agents from the wall's observing side. So Tooley has failed to supply a scenario in which causal processes may occur in opposite directions. (Ben-Yami 2007: 454)

As for Tooley, so for Dummett: whilst Dummett has stipulated that the observers in his world are unable to interfere in what they observe, those who advance the bilking argument can insist that *if* the observers could interfere, they could bilk—and so the causal processes they observe do not occur backwards, regardless of whether or not the observers interfere. It seems, then, that in creating scenarios that aim to convince the reader that participants in those scenarios are justified in concluding that the causal processes they observe occur backwards, mere stipulation that those participants are unable to interfere in the processes they observe is insufficient to thwart the bilking argument.

How might we construct a bilk-proof backwards causation scenario? We can begin by noting that the Achilles' heel of the bilking argument is the assumption that the success or failure of bilking attempts can best be explained with reference to forwards causation. This is effectively a bald, undefended denial that causation occurs backwards. The advocate of the bilking argument might justify this assumption as follows. Any scenario that involves apparent backwards causation must also involve forwards causation, since there must be something relative to which apparent backwards causation apparently runs backwards. Given this, the assumption that bilking successes and failures are best explained in terms of forwards causation is

more parsimonious, and therefore more plausible, than the assumption that they are best explained in terms of backwards causation. Were we able to devise a scenario in which the former assumption turns out to be less plausible than its denial, however, we would be on our way to a bilk-proof backwards causation scenario.

With these reflections in mind, consider a modified version of Tooley's divided world, in which a wall separates two regions where, *ex hypothesi*, causal processes occur in opposite directions. Suppose that, as in Tooley's scenario, the wall has always existed, that it is insurmountable, and that people live on both sides of the wall. Unlike in Tooley's scenario, however, light waves (but nothing else) can pass through the wall in *both* directions, and the wall is not indestructible, although nothing has yet succeeded in destroying it. The people living on each side of the wall have observed the happenings on the other side, and almost all of them have concluded that causation on the other side of the wall runs backwards relative to that on their own side. Among the people living on each side of the wall, however, are sceptics who believe that causation on the other side does not run backwards, and that if only they could break through the wall, they could prove this by bilking the alleged later causes of earlier events. Suppose that, one day, an unlikely coincidence of forces from both sides of the wall destroys a section of it. On one side of the broken section lies the golf course mentioned by Tooley. On the other side lies a cricket pitch. The group of golfers includes a sceptic about backwards causation, as does the group of cricketers. Both witness the destruction of the wall and, delighted finally to have the opportunity to lay to rest claims that causation on the other side of the wall runs backwards, each sets off to bilk the alleged later causes of earlier events.

What happens next? Let us suppose that moving to the other side of the wall does not change the direction of one's bodily (including mental) causal processes, so that when people move from the 'home' side to the 'away' side, causation on the away side continues to appear to run backwards from their point of view. As a result, anyone who moves to the away side to attempt to bilk *L*-events having observed *E*-events is not prevented from doing so by the causal limitations of their own bodily processes. This is precisely what the advocate of the bilking argument would assume, since a change in the direction of bodily causal processes would involve backwards causation, which advocates of the argument deny. With this in mind, can we suppose that the sceptics' bilking attempts will be successful?

I do not think so. Recall that by ‘backwards causation’ we do not mean causation that runs backwards relative to the temporal order, defined independently of causal considerations. Backwards causation is simply causation that runs backwards relative to some other causation. In the scenario just described, the inhabitants of each side consider causation on their side to run forwards, for the same reasons that we consider causation in our world to run forwards. Neither side’s causation is privileged by running forwards with respect to some independently determined temporal order: there is no forwards causation *simpliciter*. This means that if we wish to claim that the sceptic on one side is able in principle to bilk *L*-events on the away side, then we must concede that the sceptic on the other side can do so too: nothing about the scenario supports the claim that only one sceptic has this ability. But, both sceptics cannot have this ability. According to the bilking argument, the ability to bilk proves that all causation is forwards causation, and that any apparent backwards causation is bogus. However, since what is apparent backwards causation from the point of view of each sceptic is genuine forwards causation from the point of view of the other, either sceptic’s proving that the apparent backwards causation is bogus is incompatible with the other’s proving this. Since both sceptics cannot have the ability to bilk, and since there is no reason to suppose that either sceptic (but not the other) has this ability, the only remaining conclusion seems to be that neither sceptic has this ability. Causation really does occur in both directions in such a world.

The usual response available to the advocate of the bilking argument when presented with the systematic failure of attempts to bilk *L*-events having observed *E*-events is to claim that the failed attempts are themselves (forwards) caused by *E*-events. However, in this case, such a response will not support a denial that backwards causation occurs. Even if we concede that the bilking failures of each sceptic are caused by *E*-events, the causation from *E*-events to the bilking failures of one sceptic runs backwards relative to the causation from *E*-events to the bilking failures of the other sceptic, and so we remain committed to the view that backwards causation occurs.

Perhaps the advocate of the bilking argument can retreat to a weaker conclusion at this point. Even if backwards causation really does occur in such a world, he may insist, those who live there would not be justified in believing it to occur. Faced with their systematic failure to bilk *L*-events having observed *E*-events, sceptics from each side can choose between, on the one hand, conceding that

causation on the away side really does occur backwards, in which case they are left unable to explain the bizarre coincidences of events that conspire to prevent their bilking attempts; and on the other hand, taking *E*-events to cause the later bilking failures, that is, denying that backwards causation occurs. Choosing the latter option would allow the sceptics to dispense with the need to explain why their bilking attempts always fail for seemingly mysterious reasons—although, as Dummett noted, it leaves them unable to explain the events that appear to happen backwards on the away side. On balance, though, choosing the latter option is preferable.

However, whilst it may initially strike the sceptics that subscribing to the view that causation occurs backwards on the away side leaves them without an explanation for their bilking failures, I think that there are strong reasons to believe that they will soon realise that they already have an adequate explanation. Consider that those who observe failed attempts by sceptics from the away side to bilk the causes of events on the home side will find the failures perfectly explicable. From the point of view of the cricketers, a sceptic from the away side climbs in reverse through the hole in the wall, walks backwards onto the cricket pitch wearing an expression of frustration, tries and fails (in reverse) to prevent the bowler from bowling the ball, then runs backwards back to his own side just before the batsman hits ball at the other end of the strip. The golfers on the other side observe something analogous. Neither the observing cricketers nor the observing golfers will be at a loss to explain the failure of the bilking attempts: as with all failed attempts, failed bilking attempts will fail for a variety of reasons. The only thing the bilking attempts have in common is that they are failures; that is, they always occur just before the ball moves, a coincidence which can itself be explained with reference to the fact that the sceptic from the away side saw the ball move before he tried to bilk, and was motivated by the former to do the latter. It seems likely that sceptics who observe such happenings will come to see their own bilking failures as no more mysterious than those of the sceptics from the away side, and will realise that in seeking to explain what happens to them when they visit the away side, they need to look to their future, and not to their past. As such, their systematic failure to bilk *L*-events on the away side will come to seem no more anomalous than their systematic failure to outrun their own shadow, or to look at the back of their own head in a mirror.

As a result, the sceptics' main reason to resist conceding that causation occurs backwards on the away side disappears. Since conceding this also has the advantage

of enabling them to explain the events on the away side, concluding that causation occurs backwards on the away side becomes the most attractive option. Therefore, not only is it the case that causation in this world really does occur in both directions, it is also the case that the inhabitants of this world can justifiably believe that it occurs in both directions.⁴

FHI, Faculty of Philosophy, University of Oxford

16-

rebecca.roache@philosophy.ox.ac.uk

References

- Ben-Yami, H. 2007: 'The impossibility of backwards causation'. *The Philosophical Quarterly*, 57, pp. 439-55.
- Black, M. 1956: 'Why cannot an effect precede its cause?' *Analysis*, 16, pp. 49-58.
- Dummett, M. 1964: 'Bringing about the past'. *Philosophical Review*, 73, pp. 338-59, repr. in R. Le Poidevin and M. MacBeath (eds.) 1993: *The Philosophy of Time*. Oxford: Oxford University Press, pp. 117-33.
- Flew, A.N. 1956: 'Effects before their causes?—Addenda and corrigenda'. *Analysis*, 16, pp. 104-10.
- Flew, A.N. 1957: 'Causal disorder again'. *Analysis*, 17, pp. 81-6.
- Flew, A.N. 1964: 'Can an effect precede its cause?'. *Proceedings of the Aristotelian Society*, Supp. Vol. 28, pp. 45-62.
- Ismael, J. 2003: 'Closed causal loops and the bilking argument'. *Synthese*, 136, pp. 305-20.
- Mellor, D.H. 1998: *Real Time II*. London: Routledge.
- Reichenbach, H. 1928: *Philosophie der Raum-Zeit Lehre*, tr. as 1958: *The Philosophy of Space and Time*. New York: Dover.
- Tooley, M. 1997: *Time, Tense, and Causation*. Oxford: Clarendon Press.

⁴ I am very grateful to Robin Le Poidevin for comments on an earlier version of this paper.