Max Black and Backwards Causation

Brian Garrett

The Australian National University

Abstract

In this discussion I point out that Max Black offers not one but two arguments against the (logical/metaphysical) possibility of backwards causation. Although both arguments fail in their intended aim, they show something of importance, *viz.*, that defenders of backwards causation should understand Black's Houdini example (and others like it) in terms of the 'multiple causes' model.

Keywords: Max Black, Bilking argument, Backwards causation, Causal pre-emption, Causation.

In order to understand and assess Max Black's arguments against the possibility of backwards causation we should re-visit Black's original Houdini thought-experiment (Black 1956: 52-55). Although the so-called 'bilking argument' against backwards causation has been much discussed over the years, commentators have overlooked the fact that Black offers not one but two arguments against the possibility of backwards causation.[1] Moreover, properly formulated, Black's arguments teach us something important about the strange and complex causal laws which would have to hold in the kind of situations Black envisages.

Black's example involves an alleged case of precognition, in which Houdini announces whether a penny coin will land heads or tails. The coin is tossed one minute after his announcement. In repeated trials, Houdini always gives the right answer. Following Black's terminology we can label Houdini's predictions 'A-type events' and the outcomes of the coin tosses 'T-type events'. Let us suppose that the conditions which have obtained to date are maximally conducive to describing the example as a case of backwards causation. That is, suppose—thus far and prior to any intervention—we have good evidence for the following: Houdini's predictions exercise no causal influence on the coin tosses; the predictions and the coin tosses are not joint effects of a common cause; and no prediction was caused by a prior or simultaneous event.

Might this be considered a case of backwards causation? Black writes:

---

[1] Neither 'bilk' nor 'bilking' (nor any cognate words) appear in Black's article. The term was first introduced into the philosophical lexicon by Antony Flew two years earlier (Flew 1954: 57).

There is at least one good reason for saying "No". In the circumstances described, we can wait until *A* has happened and *then prevent* T. For we assumed that the causes of T (the way the penny was thrown, the air currents in the room, and so on) were causally independent of *A*. So, *A* having occurred, it is within our power to prevent T's occurring—by simply not tossing a coin.

This being so, a man who insists upon describing the circumstances as a case of effect preceding cause must qualify his assertion that the character of T causes *A* by adding that this is so only *when T* occurs. In other words: *T* causes *A if T* occurs, yet *A* may occur without *T* occurring at all.[2] So *A* may sometimes occur without being caused either by an earlier or a later event,[3] and yet on other occasions, a precisely similar *A,* though again not caused by an earlier event, is now held to be caused by a later one. This would be hard, if not impossible, to reconcile with our present uses of causal terminology.

But there is a still better reason for refusing to call *T* a cause of *A.* If T's causal antecedents are independent of *A* as we found ourselves required to stipulate, we can arrange for T to *disagree* with *A.* Thus, it would be theoretically possible to learn to toss the coin so that it came down heads or tails as we pleased; all we need to do is to wait for Houdini's answer, and then arrange for the coin to fall contrary to his prediction. If we *can* do this, the stipulations for the supposed precognition are logically impossible of fulfilment, because Houdini's answer will *not* always agree with the subsequent trial. On the other hand, if we find that once Houdini has answered we cannot arrange for the penny to come down as we please, we shall be compelled to say that the causal antecedents of *T* are not independent of *A.* We shall have to say that Houdini's answer exerts a causal influence of an esoteric sort upon the subsequent toss of the coin. This would be extraordinary enough, to be sure. Yet it would not be preposterous, as would be the logical absurdity of saying that an effect precedes its cause (Black 1956: 54-55).

Black offers two arguments against the possibility of backwards causation. His first argument (contained in the initial two paragraphs of the quotation) proceeds as follows:

(i) Suppose, for *reductio*, that the coin tosses cause the earlier predictions;

(ii) After a prediction has occurred, it is within our power to prevent a coin toss from occurring one minute later;

(iii) Occasionally we do so intervene (i.e., bilk); so

(iv) Houdini's predictions are sometimes caused by coin tosses (when we don't intervene) and sometimes not caused by coin tosses (when we do intervene);

(v) This conjunction (i.e., (iv)) is impossible; so

(vi) (i) is false: the coin tosses never cause the predictions; so

(vii) Backwards causation is impossible.

---

[2] Black's discussion in this paragraph and elsewhere is compromised by his failure to distinguish event-tokens from event-types. However, I assume that this can easily be rectified. What Black finds absurd is the idea that A-type events are sometimes caused by T-type events and sometimes not.

[3] Clearly this doesn't follow. If an A-type event is not caused by a later T-type event, it doesn't follow that it wasn't caused by an earlier or later event (i.e., that it was either uncaused or caused by a simultaneous event). It might have been caused by another later event. That is the main point of the present paper.

It might be questioned whether Black does endorse premise (v). He says that (iv) would be "… hard, if not impossible, to reconcile with our present uses of causal terminology" (Black 1956: 54). However Black needs (v) if he is to establish (vi) and (vii). Any weaker claim—e.g., that (iv) is merely odd or unusual—would not secure the desired result. I also assume, given his appeal to our 'causal terminology', that Black would regard (iv) as analytically impossible. I urge below that it is neither analytically nor metaphysically impossible.

Black's first argument faces the following objection. According to Black, it is absurd to suppose that Houdini's pronouncements are sometimes caused by coin tosses and sometimes not. But why is this absurd? It is not impossible for events of the same type to have different causes on different occasions. For example, suppose that switch S is normally down and, when down, causes a light to be on. But on the occasions when S is up, switch S* kicks in and causes the light to be on. The light's being on has different causes at different times. There is nothing mysterious here, and this or relevantly similar cases have doubtless occurred many times. It is, of course, a paradigm case of what is known as (forwards) causal pre-emption.

A defender of the possibility of backwards causation is entitled to suppose that backwards causal pre-emption is also possible. It could then be proposed that, in Black's example, the act of preventing the coin toss causes the earlier prediction (a prediction which otherwise would have been caused by the coin toss). On this view, Houdini's predictions are always backwardly caused, sometimes by the coin toss and sometimes by the act of preventing the coin toss.

This description is not in tension with our "causal terminology", at least if understood on the model of 'multiple causes' (i.e., different causes on different occasions). Note that I do not have to motivate the 'multiple causes' interpretation of Black's example. It is enough that the interpretation represents a metaphysically possible scenario. To refute an impossibility claim we need only describe one coherent scenario in which the alleged impossibility obtains. No other justification is needed.

On this (hitherto unexplored) way of understanding Black's example, if we let the coin toss go ahead, it causes Houdini's prediction; if we prevent the coin toss, our act of prevention causes the prediction. Either way, Houdini's prediction is an explicable and backwardly caused event. It is either caused by a coin toss or caused by an act of preventing the coin toss. So premise (v) is false.

It may, by our standards, be a strange world in which the prevention of a coin toss causes an earlier pronouncement, and such a world may have strange causal laws. But strange is not impossible. As Hume observed, it is never an *a priori* matter whether one event causes, or does not cause, another. Why has this possibility been overlooked? Perhaps because the coin toss was presented as the only salient candidate for causing the earlier pronouncement. Perhaps because it was assumed that the prevention of a coin toss is not the kind of action that could have any backwards causal effects. Whatever the explanation, this possibility represents a natural way for a defender of backwards causation to make sense of Black's example.

There are other possibilities too. For example, what if the tosser simply decides not to toss the coin? Then his failure to toss the coin causes the earlier pronouncement. Indeed, we can suppose it to be nomically guaranteed that any act or omission which results in the non-occurrence of the coin toss causes the pronouncement. Such "disjunctive" scenarios are not impossible, however weird and

gerrymandered they may be. These scenarios are "disjunctive", not in the sense that the cause is a disjunction, but in the sense that Houdini's predictions are caused in different ways on different occasions.[4]

It may be that Black endorsed premise (v) because he assumed a simple covering law model of causation. Perhaps he reasoned: if Houdini's predictions are caused by coin tosses, they must fall under an exceptionless law according to which the predictions are always caused by coin tosses; in our imagined scenario Houdini's predictions are sometimes caused by coin tosses, sometimes not; this is not possible; so coin tosses never cause the predictions. The reply to this reasoning is that its starting premise is false. The simple covering law model is refuted by cases of causal pre-emption (backwards or forwards).

Black's second, and supposedly better, argument (contained in the last paragraph of the quotation) runs as follows:

(1) Either we can arrange for the coin to fall contrary to Houdini's prediction or we cannot;

(2) If we can, "… the stipulations for the supposed precognition are logically impossible of fulfilment, because Houdini's answer will *not* always agree with the subsequent trial";

(3) If we can't, A-type events must be among the causal antecedents of T-type events (in which case we have a strange kind of forwards causation, not backwards causation);

(4) Either way, T-type events never cause A-type events; so

(5) Backwards causation is impossible.

In order for this argument to be cogent, premises (2) and (3) must be true and each must entail (4). But there is a problem.

Premise (2), as stated, is false. If we can, but never do, arrange for the coin to fall contrary to Houdini's predictions, Houdini's answers will always agree with the subsequent coin tosses. We could instead replace premise (2) with:

(2') If we can *and sometimes do* arrange for the coin to fall contrary to Houdini's predictions, Houdini's answers will not always agree with the subsequent coin tosses.

(2') may be what Black had in mind. Unfortunately, it is trivially true and fails to imply that the coin tosses never cause Houdini's predictions. (2') is consistent with the tosses causing the predictions in cases where we don't intervene. The second argument fails too.

It should be clear from the above that Black's arguments are not hopeless confusions from which we can learn nothing. Black has highlighted a perfectly intuitive worry about backwards causation, at least in situations in which agents can and do prevent the typical cause from occurring. It is puzzling what brings about the effect in such situations. But the puzzlement arises only because Black's example has been under-described. Once we think of such scenarios along the lines of the 'multiple causes' model, the worry disappears.

---

[4] I put to one side the *outre* possibility that Houdini's pronouncements, when there is no later coin toss, are uncaused and, hence, inexplicable. This may be a metaphysically possibility, but it would require even stranger laws than those envisaged above.

I realise that some may find the 'multiple causes' model hard to take seriously. It offers an extraordinarily gerrymandered case of causation. On this model, anything which results in the non-occurrence of the coin toss, after Houdini has made his prediction, causes the prediction. But 'anything' here could mean pretty much anything: intervention by a third-party, forgetfulness on the part of the coin tosser, the death of the tosser, some natural intervention (e.g., a lightning strike), … . As two referees for this journal have pointed out, this is not a kind of pre-emptive causation that we are familiar with. It is wildly disjunctive.

Let me say three things in response. First, it does not seem to be impossible. To appeal to Hume again, no causal connexion, however bizarre, can be excluded *a priori*. Second, even in cases of forwards causation, there are no *a priori* limits on what might have brought about a certain effect had the actual cause been pre-empted. The same is true in cases of backwards causation. Third, if we do not pursue some such line, it's hard to see how Black's worry can be met. If his worry can't be answered, we must concede that there are no possible worlds containing backwards causal chains and agents (or natural happenings) which can bilk those chains.[5] I would rather not concede this.

Some commentators have thought that Black's reasoning relies on an epistemic premise, *viz.*, that the intervenor knows of Houdini's pronouncement and, as a result, prevents the occurrence of the coin toss (or determines its outcome).[6] This is undoubtedly the picture that Black had in mind. But it is not essential to his arguments. The possibility of an unwitting intervenor (ignorant of Houdini and his pronouncements), or even the possibility of a natural disruptor (such as a falling tree), would serve Black's purpose just as well. Black's arguments are metaphysical not epistemic. Hence, insistence on the intervenor's ignorance of Houdini's announcements, besides being *ad hoc*, affords no resolution of Black's worry.

In sum, then, we have good grounds to reject Black's arguments. Black's opponents can render explicable Houdini's predictions if they take the acts of prevention, when they occur, to cause the earlier predictions, predictions which would otherwise have been caused by the coin tosses.[7]

References

Black, M. 1956, "Why Cannot an Effect Precede Its Cause?", *Analysis*, 16, 3, 49-58.

Flew, A. 1954, "Can an Effect Precede Its Cause?", *Proceedings of the Aristotelian Society*, 28, 45-62.

Price, H. 1996, *Time's Arrow and Archimedes' Point*, New York: Oxford University Press.

[5] This would, e.g., rule out as impossible the time travel stories beloved of philosophers.
[6] See, e.g., Price 1996: 171-74.
[7] I am grateful to two referees for this journal, and to John Broome, Alan Hajek and Daniel Stoljar for helpful feedback.