Normal Causes for Normal Effects

Totte Harinen
King's College London
This is a draft–please do not cite without permission

Abstract

Halpern and Hitchcock (2014) have used normality considerations in order to provide an analysis of actual causation. Their methodology is that of taking a set of causal scenarios and showing how their account of actual causation accords with typical judgments about those scenarios. Consequently, Halpern and Hitchcock have recently demonstrated that their theory deals with an impressive number of traditional problem cases discussed in the literature. However, in this paper I first show that the way in which Halpern and Hitchcock rule out certain cases of bogus prevention leaves their account susceptible to counterexamples. I then analyze those counterexamples in the light of the observation that, in addition to abnormal causes, people naturally focus on abnormal effects. I then argue that this bias towards abnormal effects has resulted in accounts of actual causation that cannot deal with causal scenarios in which the actual effect is normal.

1 Introduction

Following Hart and Honoré's seminal (1985) work, there have been several recent attempts to account for folk intuitions about actual causes in terms of normality (cf. Kahneman and Miller 1986, Menzies 2007, Hall 2007, Hitchcock 2007, Halpern and Hitchcock 2014). The methodology in these discussions has been that of taking a set of causal scenarios and trying to provide an account of actual causation that would agree with folk judgments about those cases. Legitimate worries can be raised as to the usefulness of that approach (cf. Glymour et al 2010, Danks 2013). But in this essay I want to put those objections to one side and discuss some causal scenarios that the normality-based accounts still don't get right. Specifically, I will focus on the recent account provided by Halpern and Hitchcock (2013; 2014), which is the most sophisticated normality-based treatment of actual causation to date.

The very basic idea in normality-based accounts has been this: out of the causally relevant factors in any given situation, ordinary speakers tend to assign a stronger causal status to those that are unexpected and abnormal. Thus, for example, if we are more likely to say that the occurrence of a forest fire is the result of someone's lighting a match rather than the result of there being oxygen present in the atmosphere–even though both factors are counterfactual difference-makers for the fire–it is because someone's lighting a match is an abnormal event whereas there

being oxygen present in the atmosphere is very normal indeed. Observations such as these have led many philosophers and psychologists to argue that actual causes are *always* exceptional events, or, as I'll often say, abnormal values of variables.

My main argument is that such accounts of actual causation suffer from 'abnormality bias'. True, in most cases, when we explain something causally, to ourselves or others, we are focusing on abnormal events. Why did the forest fire? Why did someone become pregnant? Why did that same person develop thrombosis? Why did the chief executive die? And so on. This bias is perfectly understandable because it is pragmatically important to be able to predict and control such occurrences. They provoke our natural curiosity. But such narrow focus is harmful if our goal is to give a general account of actual causation. Even if abnormal events are the most interesting kinds of effect, it doesn't follow that normal events do not have actual causes. If you build your model of actual causation by focusing on abnormal effects, there is a high risk that you cannot account for normal ones. I will demonstrate that this is exactly what has happened: Halpern and Hitchcock as well as the other advocates of normality-based accounts have only managed to give an account of the actual causes of *abnormal* effects.

The structure of my paper is the following. Because my focus is on Halpern and Hitchcock's sophisticated theory, I start by introducing the Halpern-Pearl formalism for actual causes that they use, which has been presented in various places over the last ten years (cf. Halpern and Hitchcock 2010, 2014; Halpern and Pearl 2005; Halpern 2013). This is what I will do in the next section. After that, in the following section, I will proceed to discuss several examples that Halpern and Hitchcock's account do and do not get right. In particular I will argue that their account still doesn't get certain *prevention* scenarios right. The reason for this is that those are exactly the kinds of cases in which the value of the effect variable is *normal*. This is demonstrated with examples. Because my thesis will not be fully established until I have said more about how the normality of the value of an effect variable is determined, I close my paper with a discussion on this issue. In doing so, I will draw on Tim Maudlin's (2007) ideas on causal systems obeying quasi-Newtonian laws.

2 The Halpern-Pearl Formalism

We start with a causal model M, which is a tuple (U, V, E), where U is a set of exogenous variables, V is a set of endogenous variables, and E is a set of structural equations relating the values of the variables. M is usually assumed to be acyclic, which intuitively means that there are no feedback loops in the model, and more formally that there is some total ordering $\prec$ such that, if X $\prec$ Y, then the value of X can affect the value of Y but not *vice versa*.

In the definition of actual causation below, ϕ denotes a Boolean combination of 'primitive events' of the form X = x where X $\in$ V and x is among the possible values of X. Setting the value of X in this way results in a new causal model which is otherwise identical to M except that the structural equation for X is replaced with X = x. The abbreviation $[\vec{Y} \leftarrow \vec{y}\,]$ denotes the case in which variables Y_1,..., Y_k $\in$ V are set to some of their possible values. $[\vec{Y} \leftarrow \vec{y}\,]\phi$ states that ϕ holds if Y_i

were to be set to values y_i for i = 1, ..., k. Such a 'causal formula', abbreviated ψ, requires some context $\vec{u}$. The fact that some causal formula holds in some causal model and context can then be expressed as (M, $\vec{u}$) $\models \psi$. For a more technical exposition, see any of the papers by Halpern cited in the introduction.

We can now state the Halpern-Pearl definition of actual causation (quoted from Halpern 2013, p.3 with inconsequential changes in notation and original italics):

> $\vec{X}=\vec{x}$ is an *actual cause* of ϕ in (M, $\vec{u}$) if the following three conditions hold:
>
> AC1. (M, $\vec{u}$) $\models (\vec{X}=\vec{x})$ and (M, $\vec{u}$) $\models \phi$
>
> AC2. There is a partition of V (the set of endogenous variables) into two subsets $\vec{Z}$ and $\vec{W}$ with $\vec{X} \subseteq \vec{Z}$ and a setting $\vec{x}'$ and $\vec{w}$ of the variables in $\vec{X}$ and $\vec{W}$, respectively, such that if (M, $\vec{u}$) $\models$ Z = z* for all Z $\in \vec{Z}$, then both of the following conditions hold:
>
> (a) (M, $\vec{u}$) $\models [\vec{X} \leftarrow \vec{x'}, \vec{W} \leftarrow \vec{w}] \sim \phi$
>
> (b) (M, $\vec{u}$) $\models [\vec{X} \leftarrow \vec{x}; \vec{W'} \leftarrow \vec{w}; \vec{Z'} \leftarrow \vec{z*}]\phi$ for all subsets $\vec{W'}$ of $\vec{W}$ and all subsets $\vec{Z'}$ of $\vec{Z}$, where I abuse notation and write $\vec{W'} \leftarrow \vec{w}$ to denote the assignment where the variables in $\vec{W'}$ get the same values as they would in the assignment $\vec{W} \leftarrow \vec{w}$.
>
> AC3. $\vec{X}$ is minimal; no subset of $\vec{X}$ satisfies conditions AC1 and AC2.
>
> The tuple $(\vec{W}, \vec{w}, \vec{x'})$ is said to be a *witness* to the fact that $\vec{X} = \vec{x}$ is a cause of ϕ.

Condition AC1 requires that $\vec{X} = \vec{x}$ and ϕ actually be the case. This is very reasonable, given that we are here trying to give an account of actual causation. Condition AC3 is the familiar requirement that the cause shouldn't involve elements that are redundant with respect to the effect.

Condition AC2a states that changing the value of $\vec{X}$ to some value $\vec{x}' \neq \vec{x}$ makes it the case that $\sim \phi$. This betrays the fact that the account of causation in play is counterfactual. However, it needn't be the case that setting the value of $\vec{X}$ to $\vec{x}'$ *invariably* results in $\sim \phi$. Rather, we have the partition of the endogenous variables into two sets $\vec{Z}$ and $\vec{W}$ such that $\vec{X}$ is a subset of the set of variables $\vec{Z}$. Setting the value of $\vec{X}$ to $\vec{x}'$ need only result in $\sim \phi$ under the *contingency* that the variables in $\vec{W}$ are fixed to some values $\vec{w}$. This captures the commonsensical idea that a causal dependency between two or more variables sometimes becomes 'visible' just in the case some 'off-path' variables in the causal model are kept fixed.

Condition AC2b imposes some constraints on the variables in the two sets. The requirement is that ϕ holds if we set $\vec{X}$ to $\vec{x}$ and some subset of the variables in $\vec{W}$ get the same values as they would in $\vec{W} \leftarrow \vec{w}$ and some subset of the variables in $\vec{Z}$ get the same values as they do in the actual world. The point of this requirement is to ensure that even if fixing the values of the variables in $\vec{W'}$ changes the values of the variables in $\vec{Z}$, this change is not what makes it the case that ϕ when $\vec{X} = \vec{x}$. To capture this, the relationship between $\vec{X} = \vec{x}$ and ϕ is required to hold even

when a subset of the variables in $\vec{Z}$ are set in their actual values, i.e. the values implied by (M, $\vec{u}$).

The 'witness' for the causal relation between $\vec{X} = \vec{x}$ and ϕ is the setting of variables in the relevant model that we consult in order to reveal the counterfactual dependency between the two. Importantly, such witnesses can be ranked in terms of their normality. Halpern and Hitchcock (2014) utilize a partial preorder $\succeq$ to do so. This relation is transitive and reflexive but not total.

Take two 'small worlds' s and s', which are complete specifications of the *endogenous* variables in a causal model. If s is strictly more normal than s', then Halpern and Hitchcock write s $\succeq$ s'. If s $\succeq$ s' and s' $\succeq$ s, then they write s $\equiv$ s'. In order for s and s' to be comparable, there must be some nonempty set $\vec{X}$ of variables that take more typical values in one of the worlds and less typical values in the other. However, if it is the case that there are two sets of variables $\vec{X}$ and $\vec{Y}$ such that the variables in $\vec{X}$ take more typical values in one of the worlds but the variables in $\vec{Y}$ take more typical values in the other, then s and s' are incomparable. This is what is meant by saying that $\succeq$ is not total.

Let $s_{\vec{u}}$ denote the actual small world and $s_{\vec{w},\vec{x}',\vec{u}}$ the world in which $\vec{W}$ is set to $\vec{w}$, $\vec{X}$ is set to $\vec{x}$ and the context is $\vec{u}$. Then Halpern and Hitchcock require that, in order for $\vec{X} = \vec{x}$ to be an actual cause of ϕ, conditions AC1 thru AC3 must be satisfied, and additionally in AC2a it must be the case that $s_{\vec{w},\vec{x}',\vec{u}} \succeq s_{\vec{u}}$. In other words, the small world that is the witness for the causal relation between $\vec{X} = \vec{x}$ and ϕ must be strictly more normal than the actual world. If there are many such witnesses, it is possible to order them according to their normality. The 'best witness' for $\vec{X} = \vec{x}$ being a cause of ϕ is the one that is the most normal. Halpern and Hitchcock conjecture that, in situations where there are multiple causes for some ϕ, people tend to pick out those that have the best witnesses. In what follows, whenever I talk about witnesses, I talk about *small witness worlds*.

What is the point of this normality ranking? Philosophers and psychologists have demonstrated that conceptions about what is normal play an important role in causal judgments (cf. Menzies 2007; Hart and Honoré 1985; Hitchcock and Knobe 2009; Kahneman and Miller 1986). Utilizing normality considerations together with their accounts of actual causation, philosophers have in recent years claimed to be able to explain the intuitions that ordinary speakers have about old causal chestnuts such as Prevention, Double Prevention, Bogus Prevention, Pre-emption, Omission, Selection, Switching, Emphasis, Transitivity and Thrombosis. Some of these are cases with which simple counterfactual accounts of causation struggle because the counterfactual relationship between two variables is usually masked in one way or another. This is why we need to consider a contingency–a setting of the off-path variables in our model–that reveals those masked counterfactual relationships. This is what definitions of actual causation such as the H-P account described above deliver. But even that isn't enough: we are still left with some cases where those accounts of actual causation give the intuitively wrong result. These are the cases, finally, that we hope to be able to get right by adding the normality ranking into our theory of actual causation.

3 Examples

In this section, I will discuss some causal scenarios with which Halpern and Hitchcock's normality-based account struggles. I want to emphasize that there are also several cases, which I have no space to discuss, that their account handles elegantly (for which see Halpern and Hitchcock 2014). What is common between those cases is that the actual value of the effect variable is abnormal. That is why Halpern and Hitchcock's account manages to pick out those effects' intuitively correct, abnormal actual causes. Conversely, the problematic examples that I will discuss in this section have the common feature that the value of the effect variable is normal. But, so as not to skip ahead, let me explain how Halpern and Hitchcock's account works by discussing the following perfectly ordinary, unproblematic case of pre-emption:

> EXAMPLE 1. Suzy and Bill, both competent rock throwers, are determined to destroy a nearby glass bottle. Suzy throws a rock that hits the bottle and the bottle shatters. Had Suzy not thrown her rock, Bill would have thrown his, in which case the bottle would also have shattered.

Let us model the case with the following variables:

> ST = 1 if Suzy throws, 0 otherwise
>
> BT = 1 if Bill throws, 0 otherwise
>
> BS = 1 if the bottle shatters, 0 otherwise

With these variables at hand, we can observe that the following equations capture the causal structure of the situation:

> ST = 1
>
> BT = 1 - ST
>
> BS = max (ST, BT)

The H-P definition of actual causation yields the correct result with $\vec{Z} = \{\text{ST, BS}\}$, $\vec{W} = \{\text{BT}\}$ and the value of the variable in $\vec{W}$ set to 0. The witness for the causal relation between the values of ST and BS is the small world (ST = 0, BT = 0, BS = 0). This is the contingency that we need to consider in order to unmask the counterfactual relationship between ST = 1 and BS = 1. Because that relationship is revealed only when we consider a contingency, it is easy to see that we cannot naively require that the counterfactual relationship between the values of ST and BS must be visible under all conditions in order for there to be causation between them. Finally, note that the contingency that we consider in order to unmask the relationship, namely (ST = 0, BT = 0, BS = 0), is intuitively more normal than the actual world (ST = 1, BT = 0, BS = 1) is. Thus, ST = 1 counts as an actual cause of BS = 1 under Halpern and Hitchcock's account. This is to be expected, given that the actual value of the effect variable in this causal scenario is abnormal.

Moving on to more challenging examples, let us consider a case of bogus prevention:

EXAMPLE 2. Assassin plans to put a lethal dose of poison in Victim's coffee. At the very last moment, she changes her mind and refrains. Bodyguard puts effective antidote in the coffee. Victim drinks the coffee and survives.

Here we can use the following variables:

A = (1 if Assassin puts in poison, 0 if she doesn't)

B = (1 if Bodyguard puts in antidote, 0 if she doesn't)

VS = (1 if Victim survives, 0 if she doesn't)

We need just one equation to capture the causal structure (or so I claim):

VS = max((1-A), B)

Without normality considerations, the H-P definition counts B = 1 as a cause of VS = 1 in the actual world where A = 0. For if we partition the variables into $\vec{W} = \{A\}$ and $\vec{Z} = \{B, VS\}$, set A = 1 and switch B = 0, we end up with VS = 0. AC2b doesn't help because the members of $\vec{Z}$ - $\vec{X}$ consist of $\{\}$ and VS. But observe now that the world $s_{\vec{w},\vec{x}',\vec{u}}$ is (A = 1, B = 0, VS = 0) while the world $s_{\vec{u}}$ is (A = 0, B = 1, VS = 1). Assuming that the normal values for the cause variables here are A = 0 and B = 0, we can now observe that the two worlds are, for Halpern and Hitchcock, *incomparable*. There is a singleton set $\{A\}$ with a more normal value in $s_{\vec{u}}$; and there is the singleton set $\{B\}$ with a more normal value in $s_{\vec{w},\vec{x}',\vec{u}}$. Given Halpern and Hitchcock's requirement that $s_{\vec{w},\vec{x}',\vec{u}}$ should be strictly more normal than $s_{\vec{u}}$, B = 1 does not count as a cause of VS = 1, just as required.

Let me here mark an ambiguity in Halpern and Hitchcock's account. For it is not clear whether one should incorporate the value of the *effect* variable in the normality comparison. Certainly, when Halpern and Hitchcock (2014) discuss the several examples that they are able to sort out with their account, they focus on the normality of the cause variables. A world counts as more normal than another if there is a non-empty set of variables that take more normal values in that world, and no variables that take more normal values in the other world. Elsewhere (2013, p. 17–18) Halpern and Hitchcock suggest that the normal value of the effect variable is determined by the values of the cause variables. If so, then the value of the effect variable can be ignored in the comparison simply because it is always normal. I agree that the normality of the value of the effect variable depends in some ways on the values of the cause variables, although in my view there are causal scenarios, such as prevention, in which it is nevertheless regarded as abnormal. I will return to this in detail later.

Halpern and Hitchcock's solution to bogus prevention seems worrying also because it is not quite clear how normality is relevant to the example. If you are like me, you don't tend to think that the two settings of variables–the actual setting and its witness–are somehow impossible to compare. Instead of this alleged impossibility of comparison, the real problem seems to be that B = 1 does not make any difference to the causal system given its state before B taking that value. If Halpern and Hitchcock's account picks out the intuitively correct actual cause, this seems to

be a matter of coincidence rather than a strongly rooted connection between their account and the psychological mechanisms of the attribution of actual causation.

The above speculation suggests the following question: what if the overall context in a causal scenario is changed in a way that renders the bogus preventer comparably abnormal? As it turns out, it is not too difficult to invent such cases:

> EXAMPLE 3. Assassin has switched careers and is now a bartender. Victim walks in the bar and orders a mojito. For some unknown reason, Assassin refrains from putting in alcohol. Bodyguard puts in FSB's special alcohol neutralizer nevertheless. Victim drinks the mojito and remains sober.

The variables could be:

> A = (1 if Assassin puts in alcohol, 0 if she doesn't)
>
> B = (1 if Bodyguard puts in FSB's special alcohol neutralizer, 0 if she doesn't)
>
> VS = (1 if Victim remains sober, 0 if she doesn't)

The equation describing the structure is:

> VS = max((1-A), B)

The witness $s_{\vec{w},\vec{x}',\vec{u}}$ for B = 1 being an actual cause of VS = 1 is (A = 1, B = 0, VS = 0) and the actual small world $s_{\vec{u}}$ is (A = 0, B = 1, VS = 1). Intuitively, it is now the case that $s_{\vec{w},\vec{x}',\vec{u}} \succeq s_{\vec{u}}$ given that variables A and B take more normal values in $s_{\vec{w},\vec{x}',\vec{u}}$.[1] Thus, worlds $s_{\vec{w},\vec{x}',\vec{u}}$ and $s_{\vec{u}}$ count as comparable and the witness $s_{\vec{w},\vec{x}',\vec{u}}$ is more normal than the actual world $s_{\vec{u}}$. Therefore, B = 1 counts as an actual cause of VS = 1 under Halpern and Hitchcock's account. This is so even though it seems clear that Bodyguard's putting in alcohol neutralizer doesn't play any role in Victim's remaining sober if there is no alcohol in Victim's drink in the first place. Similar scenarios are very easy to invent. This suggests that the incomparability strategy is not the correct remedy to EXAMPLE 2 either, but simply happens to deliver the correct result in that case, perhaps as a side effect. Again, the real issue seems to be that B = 1 just does not make any difference to the causal system given its state before B's occurrence. The problem with Halpern and Hitchcock's account is that it is not able to capture this feature of the causal scenario.

To conclude this section, I want to consider another case with which Halpern and Hitchcock use the incomparability strategy:

> EXAMPLE 4. Hoping for a promotion, Bodyguard wants to make it look as if she has successfully foiled a serious poisoning attempt. She convinces Cooperative Assassin to put a lethal dose of poison in Victim's coffee, but only after she has first put in some antidote. This is exactly what happens. Victim drinks the coffee and survives.

[1] For the purposes of this example, we can work with Halpern and Hitchcock's (2013) assumption that the normal value of the effect variable VS is the one determined by A and B according to the relevant structural equation. In that case, the value of VS is normal in both worlds under comparison and can be therefore ignored.

We can use the following variables:

B = (1 if Bodyguard puts in antidote, 0 if she doesn't)

C = (1 if Cooperative Assassin puts in poison, 0 if she doesn't)

VS = (1 if Victim survives, 0 if she doesn't)

The equations here are:

C = B

VS = max(B, (1-C))

B = 1 here seems to count as a cause of VS = 1 with the witness (B = 0, C = 1, VS = 0). Halpern and Hitchcock write:

> Suppose we decide that the typical value for both B and C is 0 and the atypical value 1. This would seem *prima facie* reasonable This would give us a normality ordering in which the witness (B = 0, C = 1, VS = 0) is more normal than the actual world (B = 1, C = 1, VS = 1). Intuitively, the value of B is more typical in the first world, and the value of C is the same in both worlds, so the first world is more normal overall than the second. If we reason this way, the modified theory still makes B = 1 an actual cause of VS = 1. (Halpern and Hitchcock 2014, p. 39, with small changes in notation.)

They then go on to argue that the reasoning above involves a 'subtle mistake' because it doesn't take into account the way in which the value of C depends on the value of B, and how this affects the normality of C's values. When this dependence between C and B is recognized, it turns out that the way in which C takes its values can be ranked as follows, where the normality is decreasing:

C = 0 (whatever the value of B)

C = B

C = 1 (whatever the value of B)

If this ranking is correct, Halpern and Hitchcock argue, it is no more the case that the witness world (B = 0, C = 1, VS = 0) could be compared with the actual world (B = 1, C = 1, VS = 1), because the dependence between the values of C and B is more abnormal in the witness world than in the actual world. The conclusion that Halpern and Hitchcock draw is that the two worlds are in fact incomparable and thus that B = 1 does not count as an actual cause of VS = 1.

I want to make two observations here. The first one is that, as Halpern and Hitchcock themselves point out, it is not obvious whether B = 1 should be an actual cause of VS = 1 or not. It seems as if ordinary speakers can have different intuitions about this scenario, depending on what they focus. In the next section, I'll make heavy use of this point. The second observation is that dealing with EXAMPLE 4 in the suggested way naturally adds another degree of complexity to Halpern and Hitchcock's theory. Not only do we have to evaluate the normality of the values

of variables, but now we also need to evaluate the normality of the relationships between them. This claim is not in itself unreasonable, but we are definitely on a slippery slope: as the number of factors affecting the normality of the values of variables grows, one increasingly feels that the proposed account is unfalsifiable.

4 Normal Effects

Halpern and Hitchcock think that their account can be seen 'as a formalization of Kahneman and Miller's (1986) observation that we tend to consider only possibilities that result from altering atypical features of a world to make them more typical, rather than vice versa' (Halpern and Hitchcock 2014, p. 23–24). But a closer look at Kahneman and Miller's psychological theory makes it very clear that, for them, those 'atypical features' include *effects* as well as causes:

> Causal questions about particular events are generally raised only when these events are abnormal. [...] A [causal] why question indicates that a particular event is surprising and requests the explanation of an *effect*, defined as a contrast between an observation and a more normal alternative. (Kahneman and Miller 1986, p. 148, original italics.)

So while Kahneman and Miller's research seems to provide support for the idea that people treat abnormal events as actual causes, it also suggests that the kinds of events for which people provide causal explanations are abnormal. Then the question is: what if the effect is *not* abnormal? The issue is relevant because this seems to be the case in the problematic prevention cases discussed above. Surviving, for example, can be thought of as a normal state for Victim to be, and therefore it is not clear whether its actual causes have to be something abnormal, any more than the actual causes of my getting to work at the usual time need to be something abnormal. If I am late, the situation is different–at least if I am not late regularly. Indeed, it seems that in those cases in which the effect is normal, we regard its actual causes too as normal. This is so even if it is true, as Kahneman and Miller suggest it is, that our causal why questions most of the time focus on abnormal effects. That may reveal something interesting about human psychology, but it doesn't mean that normal effects don't have actual causes. Consequently, my working hypothesis will be that normal effects require normal causes.

If normal effects require normal causes, then it is not surprising that Halpern and Hitchcock's account delivers the wrong result when applied to cases in which the actual value of the effect variable is normal. It delivers the wrong result because it recognizes only abnormal values as candidate actual causes. This claim generalizes to other similar accounts which are based on the assumption that actual causes are always unexpected, surprising or abnormal.

If I would end the discussion here, I could be accused of trickery. For I have not specified in any way what it is for the value of an effect variable to be normal. Of course, it would be possible for me to reply that the normality of the value of an effect variable is determined in the same way as the normality of the value of any other variable in a causal model, and that therefore this is not a special problem

to me but to everyone who appeals to normality rankings. But, as it happens, it is possible to say something more interesting about the issue. Let us start by considering the following example:

> EXAMPLE 5. A village gets regularly flooded in the spring, causing minor damage in some of its medieval buildings. In earlier times, the villagers have treated the floods and the resulting damage as inevitable facts of life, but this spring they take precautionary measures and erect some barriers along the river banks. As it happens, the spring is exceptionally dry and there are no floods. The medieval buildings remain undamaged.

Take the following variables:

> A = (1 if village gets flooded, 0 if it doesn't)
>
> B = (1 if barriers get erected, 0 if they don't)
>
> D = (1 if medieval buildings remain undamaged, 0 if they don't)

Let us suppose that the causal structure of the situation is as follows:

> D = max((1-A), B)

At first pass, this example looks like a problem to Halpern and Hitchcock *and* to me. For Halpern and Hitchcock, it looks like a problem for all the familiar reasons. B = 1 counts as an actual cause of D = 1 with the witness (A = 1, B = 0, D = 0), which is more normal than the actual world (A = 0, B = 1, D = 1). B = 1's being an actual cause of D = 0 is, of course, intuitively incorrect. Erecting the barrier doesn't make any difference in the causal system if there is no flood in the first place. Note, however, that the occurrence of minor damage to the village's medieval buildings, i.e. D = 1, was supposed to be a *normal* event. Thus, D = 0 is abnormal. This is a problem to me because I just argued that the reason why Halpern and Hitchcock's account of actual causation struggles with prevention scenarios is that in them the value of the effect variable is normal. But here we have their account delivering the wrong result in a prevention scenario in which the value of the effect variable appears to be *abnormal*.

Another problem to my argument is that there are, of course, cases of *genuine* prevention:

> EXAMPLE 6. Assassin first puts a lethal dose of poison in Victim's coffee. Bodyguard then manages to put in antidote. Victim drinks the coffee and survives.

With the obvious variables:

> A = (1 if Assassin puts in poison, 0 if she doesn't)
>
> B = (1 if Bodyguard puts in antidote, 0 if she doesn't)
>
> VS = (1 if Victim survives, 0 if she doesn't)

And the equation:

VS = max((1-A), B)

The point is just that here we have a case in which an intuitively abnormal value of a variable (B = 1) causes an intuitively normal value of a variable (VS = 1). Is this another counterexample to my thesis 'normal causes for normal effects'?

I shall now explain away these apparent troubles of my argument. Start from EXAMPLE 5, the village scenario. Could it be the case that the value of D = 1 is normal after all? Well, there is a sense in which it *is* normal, namely in the light of there being no floods that spring. Given that there were no floods that spring, it was normal for the medieval buildings to remain undamaged. Introducing the barriers along the river banks didn't make any difference to that. In EXAMPLE 2, the value of the effect variable (Victim's surviving = 1) is normal given the value of a cause variable (Assassin's putting in poison = 0). Bodyguard's putting in antidote doesn't change this. In EXAMPLE 3, the value of the effect variable (Victim's remaining sober = 1) is also normal, given the value of a cause variable (Assassin's putting in alcohol = 0). Bodyguard's putting in alcohol neutralizer is again irrelevant. Note that the cause variables in the light of which the values of the effect variables are normal are those that are temporally prior to the putative preventers. The significance of this can be seen by considering the two remaining scenarios.

In EXAMPLE 4, the temporally prior variable corresponds to Bodyguard's putting in antidote (B = 1). It is only later that (Cooperative) Assassin puts in poison (C = 1). In the light of B = 1, it is normal for Victim to survive (VS = 1). So here we have a similar analysis as above, except that now it is the poison that doesn't make any difference in the state of the system. But consider now EXAMPLE 6, in which the poison is added first and the antidote afterwards. Here, in the light of the value of the temporally first variable (A = 1), the normal value of Victim's surviving is 0. Crucially, Bodyguard's putting in antidote (B = 1) makes VS to take a value that is *abnormal* given the value of A. In other words, given that Assassin had put in poison, Victim's not surviving was considered normal; but Bodyguard's putting in antidote made the abnormal happen and Victim survived. This is why, I submit, EXAMPLE 6 is considered as a case of genuine prevention while the others aren't.

The pattern in these various scenarios seems to be this. In genuine preventions, the actual value of the effect variable is *abnormal* given the values of the variables temporally prior to the putative preventer. In bogus preventions, the actual value of the effect variable is *normal* given the values of the variables temporally prior to the putative preventer. One might hypothesize that the underlying psychological story goes something like this. The values of the variables that come temporally first serve to set the context, generating an expectation as to the normal value of the effect variable. The causal contribution of the putative preventer is then evaluated against this expectation. This would at least explain why, in EXAMPLES 4 and 6, it seems to make a difference whether the antidote is served before or after the poison. It also follows that, in a scenario in which the background story would be just like in EXAMPLE 4, but in which the poison would be administered first and

the antidote afterwards, people would be more inclined to treat B = 1 as a cause of VS = 1 than in EXAMPLE 4.

The above fits nicely together with Tim Maudlin's (2007) observation that special scientists as well as ordinary speakers seek to taxonomize the world into entities obeying quasi-Newtonian laws. We see such entities as having inertial states or inertial motions; and those states or motions provide the backdrop against which we identify causation. Maudlin himself argues that we tend to assign a causal status to those events that disturb the inertial states or inertial motions of the entities in our taxonomy, which is why his account is similar to the abnormality-based accounts discussed in this paper. But that is a separate component in his overall theory. What is is important for our purposes is the following connection: the inertial state or inertial motion of the entity under consideration seems to determine the normality of the effect variable in our models. The value is regarded as normal if it is what we would expect given the inertial properties of the system–and otherwise it is abnormal.

If this is correct, then it is not surprising why in the examples discussed above the expected or normal values of the effect variables change as explained. The temporally prior variables suggest an inertial motion or an inertial state for the entities–individuals and processes–that we are considering. Thus, if we know that there is flood and the inertial motion of the flood-approaching-village system is such that there is going to be some damage to the village's medieval buildings, then we expect the value of D to be '1'–and if it is not, we regard the value of D as abnormal. Then again if we know that there is no flood, the way the inertial state of our flood-approaching-village system is, the expected value of D is '0'. I claim that this is how the temporally prior variables work in all of the examples discussed above: they make us attribute an inertial state or inertial motion to the system under consideration, and it is the outcome that is in harmony with the inertial properties of the system that is the expected value of the effect variable in our model.

There is no question that our assumptions about the inertial properties of causal systems could be manipulated in many other ways. I claim no unique role for temporally prior variables, even though they seem to play this role in all of the examples that I have discussed. It is also clear that the way in which we carve the world into quasi-Newtonian entities is subject to many contingencies: as Maudlin observes, a given causal scenario can be usually interpreted in terms of several different quasi-Newtonian taxonomies. Far from being a vice, this is a virtue: it explains why some of our causal judgments are genuinely ambiguous. Consider again EXAMPLE 4. If our causal intuitions disagree about whether Bodyguard's putting in antidote is an actual cause of Victim's survival, I suggest it is because some of us focus on the inertial state of Victim when she has ingested the antidote (with the expected outcome VS = 1) while some of us focus on the inertial motion of the poison propagating in Victim's body (with the expected outcome VS = 0). It is only with reference to the latter system that Bodyguard's putting in antidote (B = 1) counts as an actual cause of VS = 1. This is also where both the actual cause and the actual effect are abnormal values of the variables in question.

5 Conclusion

I began by presenting Halpern and Hitchcock's sophisticated account of actual causation and demonstrated how it deals with some causal chestnuts. I then argued against their use of the incomparability strategy in ruling out cases of bogus prevention. I showed that adopting the incomparability strategy makes it easy to generate counterexamples by altering the background story of the relevant causal scenario and I demonstrated this by generating one such example. I then sought for an alternative way of dealing with the problematic cases of prevention. Here, I used Kahneman and Miller's observation that not only causes but also their effects are typically assumed to be abnormal events. I argued that normality-based accounts of actual causation consequently suffer from an abnormality bias, which makes them incapable of dealing with causal scenarios in which the value of the effect variable is normal. I then argued that all of the problematic cases for Halpern and Hitchcock are precisely ones in which the effect is normal. I wrapped up the discussion by putting forward a general framework of how the normality of an effect varies in causal scenarios, which was where I used Tim Maudlin's observation that special scientists as well as ordinary speakers seek to taxonomize the world into entities obeying quasi-Newtonian laws. Using this idea, I affirmed the intuitive judgment that in cases of genuine prevention the actual value of the effect variable is abnormal, while in cases of bogus prevention the actual value of the effect variable is normal.

References

Danks D (2013) Functions and cognitive bases for the concept of actual causation. Erkenntnis 78(1):111–128, URL http://link.springer.com/article/10.1007/s10670-013-9439-2

Glymour C, Danks D, Glymour B, Eberhardt F, Ramsey J, Scheines R, Spirtes P, Teng CM, Zhang J (2010) Actual causation: a stone soup essay. Synthese 175(2):169–192, URL http://link.springer.com/article/10.1007/s11229-009-9497-9

Hall N (2007) Structural equations and causation. Philosophical Studies 132(1):109–136, DOI 10.1007/s11098-006-9057-9, URL http://link.springer.com/article/10.1007/s11098-006-9057-9

Halpern JY (2013) Appropriate causal models and stability of causation

Halpern JY, Hitchcock C (2010) Actual causation and the art of modeling. In: Heuristics, Probability and Causality: A Tribute to Judea Pearl, College Publications, pp 383–406

Halpern JY, Hitchcock C (2013) Compact representations of extended causal models. Cognitive Science 37:6, 2013, pp 986-1010

Halpern JY, Hitchcock C (2014) Graded causation and defaults. British Journal for the Philosophy of Science

Halpern JY, Pearl J (2005) Causes and explanations: A structural-model approach. part i: Causes. The British journal for the philosophy of science 56(4):843–887

Hart HLA, Honoré T (1985) Causation in the Law. Clarendon Press Oxford

Hitchcock C (2007) Prevention, preemption, and the principle of sufficient reason. The Philosophical Review pp 495–532, URL http://www.jstor.org/stable/20446988

Hitchcock C, Knobe J (2009) Cause and Norm

Kahneman D, Miller DT (1986) Norm theory: Comparing reality to its alternatives. Psychological review 93(2):136

Maudlin T (2007) The Metaphysics Within Physics. Oxford University Press, URL http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199218219.001.0001/acprof-9780199218219

Menzies P (2007) Causation in context. Causation, physics, and the constitution of reality pp 191–223