

# INTERVENTIONISM DEFENDED

Kevin McCAIN

ABSTRACT: James Woodward's *Making Things Happen* presents the most fully developed version of a manipulability theory of causation. Although the 'interventionist' account of causation that Woodward defends in *Making Things Happen* has many admirable qualities, Michael Strevens argues that it has a fatal flaw. Strevens maintains that Woodward's interventionist account of causation renders facts about causation relative to an individual's perspective. In response to this charge, Woodward claims that although on his account  $X$  might be a relativized cause of  $Y$  relative to some perspective, this does not lead to the problematic relativity that Strevens claims. Roughly, Woodward argues this is so because if  $X$  is a relativized cause of  $Y$  with respect to some perspective, then  $X$  is a cause of  $Y$  simpliciter. So, the truth of whether  $X$  is a cause of  $Y$  is not relative to one's perspective. Strevens counters by arguing that Woodward's response fails because relativized causation is not monotonic. In this paper I argue that Strevens' argument that relativized causation is not monotonic is unsound.

KEYWORDS: causation, intervention, manipulation, James Woodward, Michael Strevens

A commonsense way of thinking of causal relationships is that they are relationships that can allow one to bring about changes through various manipulations. In other words, if  $X$  causes  $Y$ , then manipulating  $X$  in appropriate ways should lead to changes in  $Y$ . This commonsense intuition lies at the heart of manipulability theories of causation.<sup>1</sup> In order for a manipulability theory to have any hope of being an acceptable theory of causation it must provide an account what counts as the appropriate ways of manipulating  $X$ . A promising account of appropriate ways of manipulating  $X$  is James Woodward's account, which defines the appropriate ways of manipulating as 'interventions.' According to Woodward, "an intervention  $I$  on  $X$  with respect to  $Y$  will be such that  $I$  causes a change in  $X$ ,  $I$  does not cause a change in  $Y$  via some route that does not go through  $X$ , and  $I$  is exogenous in the sense of not itself having a cause that affects  $Y$  via a route that does not go through  $X$ ."<sup>2</sup> Woodward's interventionist account of causation

---

<sup>1</sup> See, for example, R.G. Collingwood, *An Essay on Metaphysics* (Oxford: Clarendon Press, 1940), Georg von Wright, *Explanation and Understanding* (Ithaca, NY: Cornell University Press, 1971), Peter Menzies and Huw Price, "Causation as a Secondary Quality," *British Journal for the Philosophy of Science* 44 (1993): 187-203, and James Woodward, *Making Things Happen* (Oxford: Oxford University Press, 2003).

<sup>2</sup> James Woodward, "Cause and Explanation in Psychiatry: An Interventionist Perspective," in *Philosophical Issues in Psychiatry: Explanation, Phenomenology and Nosology*, eds. K. Kendler

(hereafter referred to simply as ‘interventionism’) offers a powerful tool for understanding the nature of various causal relations that fits with the commonsense view of the connection between causation and manipulation.

Although Woodward’s case in favor of interventionism is persuasive, interventionism is not without its detractors.<sup>3</sup> Recently, Michael Strevens has attacked interventionism on the grounds that it introduces a problematic relativity to facts about causation.<sup>4</sup> More precisely, Strevens argues that interventionism is committed to the claim that the facts concerning whether  $X$  is a contributing cause of  $Y$  are dependent upon one’s perspective because contributing causation is defined with respect to a variable set. In response to this charge, Woodward attempts to provide a de-relativized notion of contributing causation.<sup>5</sup> Woodward’s response involves distinguishing between being represented as a contributing cause (following Strevens this will be referred to as ‘relativized causation’) and being a contributing cause simpliciter.<sup>6</sup> Woodward claims that although  $X$  is a relativized cause of  $Y$  relative to a particular variable set, it does not lead to the relativity that Strevens claims. Roughly, Woodward argues this is so because if  $X$  is a relativized cause of  $Y$  with respect to some variable set  $\mathbf{V}$ , then  $X$  is a contributing cause of  $Y$  simpliciter. Woodward’s argument rests on the assumption that relativized causation is monotonic in the sense that adding variables to  $\mathbf{V}$  will not lead to  $X$ ’s no longer being a relativized cause of  $Y$ .<sup>7</sup> Strevens concedes that if successful, Woodward’s response would provide a de-relativized notion of contributing causation.<sup>8</sup> However, Strevens argues that Woodward’s response fails because relativized causation is not monotonic in this

---

and J. Parnas (Baltimore, MD: Johns Hopkins University Press, 2008), 139. See Woodward, *Making Things Happen*, 98 for a more precise formal definition of ‘intervention.’

<sup>3</sup> Since my purpose here is to defend interventionism from a particular objection rather than provide a full-scale argument for its acceptance, the interested reader is encouraged to consult Woodward, *Making Things Happen*, for a thorough presentation and defense of interventionism.

<sup>4</sup> Michael Strevens, “Essay Review of Woodward, *Making Things Happen*,” *Philosophy and Phenomenological Research* 74 (2007): 233-49.

<sup>5</sup> James Woodward, “Response to Strevens,” *Philosophy and Phenomenological Research* 78 (2008): 193-212.

<sup>6</sup> Michael Strevens, “Comments on Woodward, *Making Things Happen*,” *Philosophy and Phenomenological Research* 78 (2008): 171-92.

<sup>7</sup> When Woodward and Strevens speak of adding variables to a variable set, they do not mean that the phenomenon ‘in the world’ that is being represented is changed in any way. Adding variables to a variable set amounts to giving a more detailed account of the phenomenon in question.

<sup>8</sup> Strevens, “Comments on Woodward.”

way. Given the assumed failure of Woodward's response, Strevens concludes that interventionism has the unacceptable consequence of entailing that causation is relative to one's perspective.

My goal in this paper is to defend interventionism by demonstrating that Strevens' argument is unsound. In section one, I lay out the relevant aspects of interventionism as defined by Woodward in *Making Things Happen (MTH)*. Additionally, I describe how Woodward modifies the notion of contributing cause in response to Strevens' initial objection. Also, I explain how this modification to the notion of contributing causation is supposed to meet Strevens' challenge of providing a de-relativized notion of contributing causation. In section two, I explicate Strevens' argument against the monotonicity of relativized causation. I also describe both the example that Strevens uses to support a key premise in his argument and the example that he uses to motivate his overall argument. In the third and final section, I argue for the falsity of a key premise in Strevens' argument. In addition, I argue that the example that Strevens presents to support this premise is problematic. What is more, in this section I also explicate why the example that Strevens uses to motivate his overall argument is flawed.

## 1. Interventionism and De-Relativized Contributing Causation

In order to appreciate Woodward's attempt to provide a de-relativized notion of contributing causation as well as the moves made in the dialectical exchange between Woodward and Strevens, it is necessary to be clear about the precise definitions of the following notions: *intervention*, *direct cause*, and *contributing cause*. To begin, the general idea of an intervention is fairly straightforward. Interventions are manipulations upon one or more variables in a system under idealized experimental conditions. An intervention on a variable  $X$  should be understood in terms of experimental manipulations of  $X$  that are well designed for determining if  $X$  causes  $Y$  in an idealized experimental setting (an experimental setting that excludes confounding influences). More precisely, " $I$ s assuming some value  $I=z_i$ , is an intervention on  $X$  with respect to  $Y$  if and only if  $I$  is an intervention variable for  $X$  with respect to  $Y$  and  $I=z_i$  is an actual cause of the value taken by  $X$ ."<sup>9</sup> Woodward explains that  $I$  is an intervention variable on  $X$  with respect to  $Y$  just in case:

- (IV)    I1.  $I$  causes  $X$
- I2.  $I$  acts as a switch for all the other variables that cause  $X$ . That is, certain values of  $I$  are such that when  $I$  attains those values,  $X$  ceases to

---

<sup>9</sup> Woodward, *Making Things Happen*, 98.

depend on the values of other variables that cause  $X$  and instead depends only on the value taken by  $I$ .

I3. Any directed path from  $I$  to  $Y$  goes through  $X$ . That is,  $I$  does not directly cause  $Y$  and is not the cause of any causes of  $Y$  that are distinct from  $X$  except, of course, for those causes of  $Y$ , if any, that are built into the  $I$ - $X$ - $Y$  connection itself: that is, except for (a) any causes of  $Y$  that are effects of  $X$  (i.e., variables that are causally between  $X$  and  $Y$ ) and (b) any causes of  $Y$  that are between  $I$  and  $X$  and have no effect on  $Y$  independently of  $X$ .

I4.  $I$  is (statistically) independent of any variable  $Z$  that causes  $Y$  and that is on a directed path that does not go through  $X$ .<sup>10</sup>

Simply put, the idea is that an intervention on  $X$  is some sort of change that an experimenter in an ideal setting can bring about in  $X$  that is such that the method of bringing about that change will directly affect only  $X$  and the method of changing  $X$  will exclusively set the value of  $X$ .

The next notion that needs to be defined is *direct cause*. According to Woodward,  $X$  is a direct cause of  $Y$  with respect to variable set  $\mathbf{V}$  if and only if there is “a possible intervention on  $X$  that will change  $Y$  (or the probability distribution of  $Y$ ) when all other variables in  $\mathbf{V}$  besides  $X$  and  $Y$  are held fixed at some value by interventions.”<sup>11</sup>

Now the notion that lies at the heart of the debate between Woodward and Strevens, *contributing cause*, needs to be examined. Woodward defines a contributing cause in the following manner:

A necessary and sufficient condition for  $X$  to be a (type-level) contributing cause of  $Y$  with respect to variable set  $\mathbf{V}$  is that (i) there be a directed path from  $X$  to  $Y$  such that each link in this path is a direct causal relationship; that is, a set of variables  $Z_1 \dots Z_n$  such that  $X$  is a direct cause of  $Z_1$ , which is in turn a direct cause of  $Z_2$ , which is a direct cause of  $\dots Z_n$ , which is a direct cause of  $Y$ , and that (ii) there be some intervention on  $X$  that will change  $Y$  when all other variables in  $\mathbf{V}$  that are not on this path are fixed at some value. If there is only one path  $P$  from  $X$  to  $Y$  or if the only alternative path from  $X$  to  $Y$  besides  $P$  contains no intermediate variables (i.e., is direct), then  $X$  is a contributing cause of  $Y$  as long as there is some intervention on  $X$  that will change the value of  $Y$ , for some values of the other variables in  $\mathbf{V}$ .<sup>12</sup>

---

<sup>10</sup> Woodward, *Making Things Happen*, 98. Woodward notes that any unqualified instance of ‘cause’ in this definition should be understood to mean contributing cause.

<sup>11</sup> Woodward, *Making Things Happen*, 55.

<sup>12</sup> Woodward, *Making Things Happen*, 59.

Strevens argues that interventionism makes causation relative to one's perspective because according to Woodward's definition, whether or not  $X$  is a contributing cause of  $Y$  is a relative matter.<sup>13</sup> He supports this claim by pointing out that contributing cause is defined with respect to a variable set. So, Strevens challenges Woodward to explain how interventionism can provide an account of contributing causation that is not dependent upon one's perspective. That is, Strevens challenges Woodward to de-relativize the notion of contributing cause.

In light of this criticism and the challenge put forward by Strevens, Woodward makes the following modifications. First, he claims that the above definition of contributing cause from *MTH* would have been better put as necessary and sufficient conditions for " $X$  to be correctly represented as a contributing cause of  $Y$  with respect to  $\mathbf{V}$ ."<sup>14</sup> Second, he claims that "One can then go on to say that  $X$  is a contributing cause of  $Y$  *simpliciter* (in a sense that isn't relativized to any particular variable set  $\mathbf{V}$ ) as long as it is true that there exists a variable set  $\mathbf{V}$  such that  $X$  is correctly represented as a contributing cause of  $Y$  with respect to  $\mathbf{V}$ ."<sup>15</sup>

Given Woodward's modifications it is easy to see how he responds to Strevens' challenge. First, the original notion of contributing cause is not to be understood as a definition of contributing causation, but instead as a definition of representation as a contributing cause (relativized causation). Second, Woodward offers a de-relativization of contributing causation by claiming that  $X$  is a contributing cause of  $Y$  (in a non-relativized sense) if and only if  $X$  is a relativized cause of  $Y$  relative to some variable set. As Strevens notes, Woodward's de-relativization of contributing causation is a success only if relativized causation is monotonic.<sup>16</sup> That is, Woodward has defended interventionism from the charge of relativity only if it is true that "if  $X$  is a relative cause of  $Y$  with respect to a variable set  $V$ , then it is also a relative cause of  $Y$  with respect to any superset of  $V$ ."<sup>17</sup>

Now that the relevant items of interventionism have been defined and Woodward's response to Strevens' challenge explained, I will explicate Strevens' argument against the monotonicity of relativized causation.

---

<sup>13</sup> Strevens, "Essay Review of Woodward."

<sup>14</sup> Woodward, "Response to Strevens," 209.

<sup>15</sup> Woodward, "Response to Strevens," 209.

<sup>16</sup> Strevens, "Comments on Woodward."

<sup>17</sup> Strevens, "Comments on Woodward," 175.

## 2. Strevens' Attack on Monotonicity

Strevens argues that Woodward's attempt to provide an unrelativized notion of contributing causation fails because relativized causation is not monotonic. More specifically, Strevens argues that  $X$  may be a relativized cause of  $Y$  relative to variable set  $\mathbf{V}$ , but not a relativized cause of  $Y$  relative to variable set  $\mathbf{V}^*$  (a variable set constructed by adding more variables to  $\mathbf{V}$ ). Strevens offers the following formulation of his argument:

1. Adding variables to a variable set can sometimes make relativized causal relations appear (as monotonicity allows).
2. A variable's counting as an intervener depends on the *non-existence* of certain relations of relativized causation.
3. Thus (from (1) and (2)), variables may lose their status as interveners as other variables are added to the variable set.
4. A variable's status as a relativized cause requires the existence of an intervener with respect to which a certain further condition is satisfied. If a variable loses its status as an intervener, then, other variables may lose their status as relativized causes.
5. Thus (from (3) and (4)), variables may lose their status as relativized causes as other variables are added to the variable set.<sup>18</sup>

Concerning line one, Strevens correctly notes that it is consistent with monotonicity. One can easily see how this premise is true given interventionism. To illustrate this, consider a variable set,  $\mathbf{V}$ , which includes only  $X$  and  $Y$  where  $X$  is a relative cause of  $Y$ . Suppose further that as a matter of fact there are intermediate causal links between  $X$  and  $Y$  (this is an unproblematic supposition because it is plausible that there are intermediate links between any causally related variables – excepting, perhaps, those that are representative of features of fundamental physical reality). If the variable set  $\mathbf{V}$  were supplemented with one of these intermediate causal links,  $Z$ , relativized causal relations will appear in the augmented variable set  $\mathbf{V}^*$ . Specifically, relativized causal relations between  $Z$  and both  $X$  and  $Y$  that did not hold relative to  $\mathbf{V}$  will hold relative to  $\mathbf{V}^*$ .<sup>19, 20</sup>

---

<sup>18</sup> Strevens, "Comments on Woodward," 175-76.

<sup>19</sup> Strevens also appeals to an example to help support this premise. However, it is not necessary to describe this example in detail because the example is problematic and, further, it is not needed to support this premise.

<sup>20</sup> Michael Strevens has mentioned (personal correspondence) that this is not quite how he intended to support this premise. Strevens thinks that this premise is true because he thinks that adding variables to a variable set,  $\mathbf{V}$ , can lead to the appearance of relativized causal relations

Line two of this argument is true largely by definition. According to Woodward, *I*'s being an intervention on *X* with respect to *Y* requires that *I* not be a contributing cause of *Y* via some causal path that does not go through *X*. Strevens points out that the notion of an intervention does not explicitly appeal to relativization to a variable set, but it does appeal to the notion of contributing causation. He correctly notes that the notion of contributing cause that is appealed to in the definition of an intervention is either relativized or not. If it is relativized, then this premise is true. If it is not relativized, then it not clear that Woodward's attempt to provide an unrelativized notion of contributing causation is successful. As Strevens says "it is far from clear that an account of unrelativized causation that takes the form of a definition invoking unrelativized causal facts constitutes a genuine derelativization."<sup>21</sup> While it is not certain that an unrelativized notion of causation cannot be crafted in this way, Strevens is correct in claiming that it is not clear that it can. So, the truth of this premise should be granted.

Strevens supports his premise in line four by way of an example. In Strevens' example an experimenter is interested in determining whether bottled water consumption is a cause of heart disease. According to Strevens, a bungling experimenter may manipulate subjects' bottled water consumption by increasing their intake of salty foods. Since eating salty foods increases one's chance of heart disease, there will be a correlation between increased bottled water consumption and heart disease in this case. Strevens claims that relative to the variable set that only includes bottled water and heart disease; salty food intake will count as an intervention on bottled water consumption with respect to heart disease. So, relative to this variable set bottled water consumption will count as a cause of heart disease. However, Strevens maintains that if we augment this variable set with other variables, such as artery hardening, a relativized causal relation between salty food intake and heart disease will appear; thus, revoking the status of salty food intake as an intervention on bottled water consumption. As a result of the failure of salty food intake to count as an intervention relative to this variable set, bottled water consumption will lose its status as a relativized cause of heart disease.

---

between the members of the original set **V**. For example, Strevens thinks that even though *X* and *Y* are causally unrelated relative to **V**, adding variables to **V** can make a relativized causal relation between *X* and *Y* appear. I think that Strevens is mistaken on this point, however, since Strevens and I both agree that this premise is true, the issue of what makes this premise is true can be set aside for the moment.

<sup>21</sup> Strevens, "Comments on Woodward," 181.

Although Strevens uses similar examples to support the premises in lines one and four of his argument, he notes that these two examples each involve a different set of assumptions. So, in an effort to bolster his overall argument Strevens presents a third example, which purports to describe a situation in which both premise one and premise four are true. Here is Strevens' third example. Again he is considering the relationship between drinking bottled water ( $B$ ) and heart disease ( $H$ ). In this example Strevens stipulates that eating salty foods ( $S$ ) sometimes leads one to drink bottled water and sometimes leads one to drink red wine ( $W$ ). He assumes that if eating salty foods leads one in a particular instance to drink bottled water, then it does not lead to her drinking red wine; and vice versa. He also stipulates that this sporadic consumption of red wine is enough to off-set the effect that eating salty foods has on heart disease by hardening one's arteries ( $A$ ). Strevens points out that eating salty foods may be used to manipulate bottled water consumption, but it will only be successful some of the time because eating salty foods sometimes leads to red wine consumption. However, Strevens claims that in the subset of cases where eating salty foods successfully manipulates bottled water consumption there will be a correlation between bottled water consumption and heart disease because there will be no red wine consumption to off-set the effect eating salty foods has on heart disease. He goes on to say that the fact that  $S$  is not a legitimate intervention on  $B$  with respect to  $H$  will only show up when the variable set under consideration includes either  $W$  or  $A$ . So, Strevens maintains that  $B$  will be a relativized cause of  $H$  relative to the variable set  $\{B, H, S\}$ . However, since  $S$  will not be an intervention on  $B$  with respect to  $H$  relative to variable sets  $\{B, H, S, W\}$ ,  $\{B, H, S, A\}$ , or  $\{B, H, S, W, A\}$ ,  $B$  will not be a relativized cause of  $H$  relative to any of these variable sets. Strevens argues that his example illustrates a situation where  $B$  is a relativized cause of  $H$  relative to a particular variable set, but  $B$  fails to be a relativized cause of  $H$  relative to variable sets constructed by adding further variables to the original variable set. Thus, Strevens concludes that this example reveals that relativized causation is not monotonic.

### 3. In Defense of Monotonicity

My strategy for defending Woodward's claim that relativized causation is monotonic from Strevens' attack is straightforward. I argue that the premise in line four of Strevens' argument, "If a variable loses its status as an intervener, then, other variables may lose their status as relativized causes," is false. First, I argue that the example that Strevens appeals to in defending this premise fails to provide evidence for its truth. Second, I explicate why Strevens' third example, which is

designed to motivate his overall argument, is problematic. Third, I explain why in general arguments of the kind Strevens offers fail to establish that relativized causation is not monotonic.

In order to demonstrate that the premise “If a variable loses its status as an intervener, then, other variables may lose their status as relativized causes” is false, it is important to first spend a little time re-examining the notion of an intervention. Specifically, it is necessary to examine the conditions under which a variable  $I$  counts as an intervention on  $X$  with respect to  $Y$ . For the sake of simplicity, in the course of evaluating these conditions I will consider a very sparse variable set,  $\mathbf{V}$ , which includes only  $I$ ,  $X$ , and  $Y$ . According to Woodward’s definition  $I$  is an intervention on  $X$  with respect to  $Y$  just in case  $I$  is an actual cause of the value taken by  $X$  and  $I$  is an intervention variable for  $X$ . In order to be an intervention variable for  $X$ ,  $I$  must satisfy four further conditions of **IV**: 1)  $I$  has to cause  $X$ , 2)  $I$  has to act as a switch for all other variables that cause  $X$ , 3)  $I$  must not be a direct cause of  $Y$  nor a cause of any causes of  $Y$  that are distinct from the causal connection  $I$ - $X$ - $Y$ , and 4)  $I$  must be statistically independent of any variables that cause  $Y$  without causing  $X$ . Utilizing the sparse variable set mentioned above  $\{I, X, Y\}$ , let us assume that  $I$  is an intervention on  $X$  with respect to  $Y$ . So,  $I$ ’s having the value that it does causes  $X$  to have its actual value and  $I$  is solely responsible for  $X$  having that value (this follows from the first condition of an intervention and the first two conditions of **IV**). Additionally, there is no possible intervention on  $I$  that will change the value of  $Y$  when  $X$  is held fixed at a certain value via other interventions (this follows from condition three of **IV** and the definition of direct cause). Finally, there are no causes of  $Y$  that are also causes of  $I$  (this follows from condition four of **IV**).

At this point I will turn to a critical assessment of the example Strevens uses to support the premise “If a variable loses its status as an intervener, then, other variables may lose their status as relativized causes.” Recall that Strevens’ example involves the incompetent experimenter who uses salty food intake ( $S$ ) to manipulate bottled water consumption ( $B$ ) in order to determine if bottled water consumption ( $B$ ) causes heart disease ( $H$ ). Strevens maintains that if we simply consider the variable set  $\{S, B, H\}$ ,  $S$  will be an intervention on  $B$  with respect to  $H$ . So,  $B$  will be a relativized cause of  $H$  relative to this variable set because using  $S$  to intervene on  $B$  will lead to a change in  $H$ . However, he claims that if further variables such as artery hardening ( $A$ ) were added to this variable set,  $S$  would lose its status as an intervention. Thus,  $B$  would not be a relativized cause of  $H$  relative to the augmented set  $\{S, B, H, A\}$ .

This example is ineffectual because, contrary to what Strevens claims, *S* is not an intervention on *B* with respect to *H* even in the impoverished three variable set. The reason that this is the case is obvious. *S* fails to meet the third condition of **IV**. That is, relative to the variable set containing  $\{S, B, H\}$  *S* will be a direct cause of *H*. *S* is a direct cause of *H* relative to this variable set because there are possible interventions on *S* that lead to changes in *H* while *B* is held fixed. Keeping with Strevens' example, one such intervention would be for the experimenter to lock the subjects in a room with only salty food to eat while providing the subjects a fixed quantity of bottled water to drink and nothing else. If the experimenter intervenes on salty food intake while holding bottled water consumption fixed in this manner, there will be a correlation between salty food and heart disease.

One might worry that this way of responding to Strevens' example sneaks in an unrelativized notion of causation into the notion of an intervention. The concern here is that whether *S* is a direct cause of *H* relative to this variable set depends on relations between *S*, *B*, *H*, and variables outside of the variable set under consideration. One might think that what is really occurring here is an illegitimate appeal to facts about all of the variables there are for determining what counts as an intervention. So, the worry is that appealing to these facts about variables outside of the variable set under consideration is utilizing unrelativized causation.<sup>22</sup>

There are two ways of responding to this worry. The first way to respond is to point out that there is nothing in this response to Strevens' example that appeals to an unrelativized notion of causation. Direct causation is a relativized notion of causation. Further, appealing to the existence of variables that can possibly be added to a variable set is not invoking unrelativized causation. The second way to respond is to draw attention to the fact that the satisfaction of the third condition of **IV** only requires that it not be possible to manipulate *S* in a way that will affect *H* while *B* is held fixed. Again nothing about this suggests an illicit appeal to unrelativized notions of causation.

---

<sup>22</sup> This objection bears some similarity to an unrelated objection that Strevens, "Essay Review of Woodward," raises for Woodward's account of causation. Strevens argues that in order to determine whether *I* is an intervention we need to know about the causal relations that *I* bears to other variables. In order to determine these relations we need to intervene on *I*, so we need another intervention variable *I*\* that is an intervention on *I*. However, in order to determine whether *I*\* is an intervention we need to know about the causal relations that *I*\* bears to other variables, and so on. Thus, there seems to be a problem with ever determining whether *I* is an intervention.

Given my treatment of the example Strevens uses to support this premise, the problem with the example he uses to motivate his overall argument is probably apparent. However, at the risk of a bit redundancy I will briefly explain what is wrong with this example as well. Remember that in this example Strevens stipulates that eating salty foods ( $S$ ) sometimes leads one to drink bottled water ( $B$ ) and sometimes leads one to drink red wine ( $W$ ). He assumes that if eating salty foods leads one in a particular instance to drink bottled water, then it does not lead to her drinking red wine; and vice versa. Further this occasional consumption of red wine is enough to off-set the effect that eating salty foods has on heart disease ( $H$ ) by hardening one's arteries ( $A$ ). Strevens claims that in the subset of cases where eating salty foods successfully manipulates bottled water consumption there will be a correlation between bottled water consumption and heart disease because there will be no red wine consumption to off-set the effect eating salty foods has on heart disease. He believes that the fact that  $S$  is not a legitimate intervention on  $B$  with respect to  $H$  will only show up when the variable set under consideration includes either  $W$  or  $A$ . So, Strevens maintains that  $B$  will be a relativized cause of  $H$  relative to the variable set  $\{B, H, S\}$ . However, since  $S$  will not be an intervention on  $B$  with respect to  $H$  relative to variable sets  $\{B, H, S, W\}$ ,  $\{B, H, S, A\}$ , or  $\{B, H, S, W, A\}$ ,  $B$  will not be a relativized cause of  $H$  relative to any of these variable sets.

This example shares the same problem as the previous example. Namely,  $S$  does not meet the conditions for being an intervention on  $B$  with respect to  $H$ . Again in this example  $S$  is a direct cause of  $H$  relative to this variable set because there are possible interventions on  $S$  that lead to changes in  $H$  while  $B$  is held fixed. One such intervention would be the one I described above, giving someone only salty food to eat and a limited supply of bottled water to drink and nothing else. In this situation changes in  $S$  would be correlated with changes in  $H$  while  $B$  is held fixed. Since  $S$  is not an intervention on  $B$  with respect to  $H$ ,  $B$  will not be a relativized cause of  $H$  relative to the variable set  $\{S, B, H\}$ . So, Strevens' example fails to illustrate a situation where a variable that is a relativized cause relative to a variable set ceases to be so when more variables are added to the variable set.

One might try to defend Strevens' example from my objection by having  $S$  represent not only the eating of salty food, but instead the eating of salty food in an environment where bottled water and red wine are both freely available. If  $S$  represents the eating of salty foods in an environment where bottled water and red wine are both freely available, one may think that the experiment that I describe above will not count as an intervention on  $S$ . The idea is that given what  $S$  represents in this case, an experiment that only gives someone salty food to eat

and bottled water to drink will not be intervening on  $S$  because the experiment requires an environment that differs from the one specified by  $S$ . So, the experiment that I describe would fail to be an intervention on  $S$  with respect to  $H$ . Since the experiment does not intervene on  $S$ , it cannot show that  $S$  is a direct cause of  $H$ . Thus, the experiment described above cannot demonstrate that  $S$  is not an intervention on  $B$  with respect to  $H$ , and so, it fails to pose a problem for Strevens' example.<sup>23</sup>

Although it is true that understanding  $S$  as representing eating salty foods in the specified environment will make it the case that the experiment I describe is not an intervention on  $S$  with respect to  $H$ , this move will not save Strevens' example. The problem for Strevens' example does not go away by having  $S$  specify a particular environment. Even if  $S$  represents eating salty foods in an environment where bottled water and red wine are both freely available, there are possible interventions on  $S$  that lead to changes in  $H$  while  $B$  is held fixed. One such intervention would be to perform an experiment where subjects are given only salty food to eat and they are forced to drink so much bottled water that although there is red wine available, they will not drink any. This experiment will count as an intervention on  $S$  and it has the same result as the original experiment that I described, namely, it results in subjects who eat salty foods and only drink bottled water. So, this experiment will expose the causal connection between  $S$  and  $H$  because changes in the amount of salty food eaten in this circumstance will lead to changes in heart disease, and hence, this experiment demonstrates that  $S$  is not an intervention on  $B$  with respect to  $H$ . Thus, even if  $S$  represents eating salty foods in a specified environment, Strevens' example is still problematic.

Now that I have shown why Strevens' argument is unsuccessful I will explain why other arguments of this kind will also fail to demonstrate that relativized causation is not monotonic. As noted above, there are very specific conditions under which  $I$  may be properly said to be an intervention on  $X$  with respect to  $Y$ . Given an understanding of these conditions, it is possible to see how adding variables to variable set  $\mathbf{V}$ , which includes  $\{I, X, Y\}$ , might result in  $I$ 's no longer meeting those conditions. Since  $I$  is the sole cause of  $X$  having the value that it does and  $I, X, Y$  are the only variables in the set,  $I$  will be a direct cause of  $X$  relative to this set. Adding variables to  $\mathbf{V}$  can make it the case that  $I$  is no longer a direct cause of  $X$ . For instance, there may be a variable  $Z$  that is a causal link between  $I$  and  $X$ . However, this fact will not mean that  $I$  is no longer an intervention because adding variables will not make it the case that  $I$  fails to be a cause of  $X$ . So, adding variables will not lead to  $I$ 's failing to satisfy the first

---

<sup>23</sup> Thanks to Michael Strevens for drawing my attention to this objection.

condition for being an intervention nor the first condition for being an intervention variable. The second condition of **IV** requires  $I$  to essentially sever the causal links between  $X$  and all other variables besides  $Y$ . It seems possible that if enough variables are added to  $\mathbf{V}$ , there will be some for which  $I$  fails to function as a switch. So, it seems possible that adding variables to  $\mathbf{V}$  may lead to  $I$  failing to meet the second condition of intervention variables. Since  $I$  is not a direct cause of  $Y$  relative to  $\mathbf{V}$ , then adding variables to the set will not lead to  $I$  becoming one. Likewise, adding variables to the set will not lead to  $I$  becoming a cause of  $Y$  along another path than the  $I$ - $X$ - $Y$  path. So, adding variables will not lead to  $I$ 's failing to satisfy condition three of **IV**. Finally, if  $I$  is statistically independent of any cause of  $Y$  (as condition four requires), then adding further variables to  $\mathbf{V}$  will not lead to  $I$  and  $Y$  having a common cause. So, it seems that the only way that adding variables to a variable set can remove  $I$ 's status as an intervention is by introducing variables that are causally related to  $X$  that are such that  $I$  cannot block their causal impact on  $X$ .

In order for arguments like Strevens' to be successful it has to be the case that there are ways that a variable can lose its status as an intervener which result in other variables losing their status as relative causes. However, the only way that  $I$  can lose its status as an intervener by adding variables to  $\mathbf{V}$  is for it to fail to disrupt the causal connections between some of these new variables and  $X$ . This situation will not result in variables losing their status as relativized causes in general and it will not result in  $X$  losing its status as a relativized cause of  $Y$  in particular. So, an evaluation of the conditions for  $I$  being an intervener shows that Strevens' argument and other arguments of the same kind cannot demonstrate that relativized causation is not monotonic.

Assuming that what I have said here is correct, Strevens fails to show that relativized causation is not monotonic. More generally, arguments of the kind that Strevens presents cannot establish that relativized causation is not monotonic. Thus, it is reasonable to conclude that Woodward's attempt to provide an unrelativized notion of contributing causation is successful. Which in turn means that it is reasonable to think that interventionism does not render causation relative in the problematic way that Strevens suggests.<sup>24</sup>

---

<sup>24</sup> Thanks to Michael Strevens and Jim Woodward for helpful discussion of this topic. A very special thanks to Brad Weslake for comments on numerous earlier drafts and for many fruitful discussions of these issues.