

An Epistemology of Causal Inference from Experiment

Karen R. Zwier

Abstract

The manipulationist account of causation provides a conceptual analysis of cause-effect relationships in terms of hypothetical experiments. It also explains why and how experiments are used for the empirical testing of causal claims. This paper attempts to apply the manipulationist account of causation to a broader range of experiments—a range that extends beyond experiments explicitly designed for the testing of causal claims. I aim to show (1) that the set of causal inferences afforded by an experiment is determined solely on the basis of contrasting case structures that I call “experimental series”, and (2) that the conditions that suffice for causal inference obtain quite commonly, even among “ordinary” experiments that are not explicitly designed for the testing of causal claims.

1. Introduction. The manipulationist account of causation, exemplified especially in the work of Woodward (2003), is a powerful and interesting explication of the meaning of causal claims. The account is intended as a conceptual clarification of what it is to be a causal relationship, and it provides this clarification by making reference to hypothetical experiments and ideal interventions. And since, according to the account, hypothetical experiments are embedded in the very content of causal claims, it requires only a small logical step to explain the role of experimentation in the empirical investigation of causal claims.

No one can deny that *some* scientists intend to test causal claims, and that they design and carry out experiments for the purpose. Does the type of fertilizer applied to potatoes affect crop yield? A scientist might perform an experiment by applying different types and quantities of fertilizer and comparing the resulting yield. Does a certain drug improve prognosis for patients with a certain condition? A group of scientists might perform a series of randomized, double-blind trials to find out. The manipulationist account certainly seems to be applicable for analyzing the success or failure of causal inference in experiments such as these. However, it is not quite as easy to see if—or how—the manipulationist account might apply to experiments that are *not* explicitly designed or carried out for the purpose of testing a causal claim.

Experiments in the physical sciences, in particular, rarely seem to be framed in terms of causal questions, at least not explicit ones. Consider an experiment aimed at measuring the boiling temperature of nitric acid at atmospheric pressure. Is such an experiment intended to test a causal claim? It certainly doesn't seem so, at least not at first glance. But could the experiment still afford causal inference, if we knew where to look and what assumptions to apply? I take the answer to this latter question to be non-obvious, and the goal of this paper is to make some progress toward an answer.

This paper attempts to apply the manipulationist account of causation to a broader range of experiments—a range that extends beyond the set of experiments that are explicitly designed for the testing of causal claims. I wish to include anything that we might naturally call an “experiment”—i.e., a scientific study in which the investigator deliberately sets up and/or intervenes on a system for the purpose of studying it.¹ I aim to show (1) that the set of causal inferences afforded by an experiment is determined solely on the basis of contrasting case structures that I call “experimental series”, and (2) that the conditions that suffice for causal inference obtain quite commonly, even among “ordinary” experiments that are not explicitly designed for the testing of causal claims.

The implications of this point are potentially far-reaching. Even experiments not branded as “causal”, including those carried out in the course of research in the physical sciences, can, under certain circumstances, afford causal inference. As a result, an experiment that meets certain criteria has the ability to furnish causal content even in those areas of science (e.g., fundamental physics) where causal content is less obvious.

2. The Manipulationist Account of Causation. I begin with a brief overview of the manipulationist account of causation. The manipulationist account, in its most basic form, is intended as an account of the *meaning* of causal claims. A meaningful causal claim must have an interpretation that refers to the result of some relevant hypothetical experiment. But what is the relevant hypothetical experiment for a given causal claim? Roughly, the idea is the following: for a causal claim such as “*X* causes *Y*”, the hypothetical experiment under

¹Purely observational studies (e.g., observing astronomical events through a telescope, analyzing retrospective health information, etc.) that involve no intervention or set-up on the part of the investigator will not be considered experiments for my purposes here.

consideration is one in which the variable or factor X is manipulated or changed in some way, and any corresponding change (or non-change) in Y is observed. According to the manipulationist account of causation, consideration of such an experiment is logically embedded in the very content of a well-formed causal claim, such that evaluation of the truth or falsity of the claim will be tied to an evaluation of whether or not a change in X would be seen if the experiment were to be performed.

We can state the idea more formally as a criterion for X to be considered a cause of Y :

MANIPULATIONIST CAUSE: X is a *cause*² of Y iff, under some set of background conditions $\mathbf{BC} = \{BC_1, BC_2, \dots, BC_n\}$ having values $\{bc_1, bc_2, \dots, bc_n\}$, given some (possibly empty) set $\mathbf{S} = \{S_1, S_2, \dots, S_m\}$ of variables other than X and Y that are held fixed at predetermined values $\{s_1, s_2, \dots, s_m\}$, there is some ideal intervention³ I on X that would change the value of Y .

²“Cause”, as I use it here and throughout this paper, corresponds to Woodward’s “type-level contributing cause”. The criterion that I give here is a modified and simplified version of Woodward’s **M**, which requires detailed knowledge of the path from X to Y (see Woodward (2003, 59)). In the context of my discussion here, I do not wish to assume that an evaluator of causal claims always has that knowledge, and so I give a criterion that does not require it. In addition, my criterion is intended to be more faithful to the implicit criterion for causation in the mind of an actual experimenter who is—implicitly or explicitly—testing a causal claim.

³The manipulationist account requires that an intervention variable have particular characteristics in relation to both X and Y and the larger system of variables being considered. For the purpose of brevity, I will not discuss these requirements here; see Woodward (2003).

According to the above criterion, a hypothetical experiment relevant to the evaluation of the claim “ X is a cause of Y ” is one in which we hold some (possibly empty) set of variables S fixed while intervening on X , and we observe any associated changes in the value of Y . The claim “ X is a cause of Y ” will be true if and only if changes would be observed in Y in the context of *some* hypothetical experiment defined by a specific BC , S , and I .

An important thing to note about this way of spelling out the meaning of a causal claim is that it makes use of a particular kind of counterfactual claim. In order to make sense of how an intervention on one variable, X , “makes a difference” to another variable, Y , we need to have some concept of what *would have happened* had the intervention on X not occurred. It is only by comparing the case in which the intervention is performed with our background understanding of what would have happened had the intervention not been performed (or had a different intervention been performed) that we get a sense of an effect.

A second thing to notice about the hypothetical experiment referenced by a causal claim is that it involves two different types of interactions with the experimental system. The values of the background condition variables in BC are observed, as is the value of Y . Nothing is done to directly force these variables to take on particular values. For X and for the set S , however, interventions directly force these variables to take on certain values.⁴ The distinction between *observing* the value of certain variables and *intervening* to set the value of others is absolutely central to the manipulationist account of causation. The character of the knowledge that we

⁴Experiments with a non-empty S will be multiple-intervention experiments intended for ruling out “unfaithfulness”, as it is called in the causal modeling literature. In cases of unfaithfulness, observational data (and even some experimental data) can make it appear that two variables are independent of one another despite one being a cause of the other. See Spirtes et al. (2000, 13–14), Woodward (2003, 49–50), and Zhang and Spirtes (2008).

gain from observing a natural course of events in a system and that of the knowledge that we can gain from carefully designed interventions on that same system are essentially different. When we know from mere observation that certain values of X are associated with certain values of Y , this fact underdetermines the various types of causal connections that might exist between the two variables. Assuming that the correlation is not a spurious result of sample or selection bias, there are three different ways in which the variables might be causally connected: (i) X could be a cause of Y , (ii) Y could be a cause of X , and/or (iii) X and Y could share a common cause (or set of common causes). Interventions allow us to distinguish among these three types of causal connections (and their several combinations), because each kind of causal connection between X and Y would respond differently to interventions on X or Y .

3. From Hypothetical Experiment to Real Experiment. The conceptual tools and criteria discussed in the previous section serve the primary goal of the manipulationist account of causation: that of explicating and interpreting causal claims in terms of hypothetical experiments. Given a causal claim, these tools allow us to reconstruct the relevant hypothetical experiment embedded in the claim (or a set of relevant hypothetical experiments that reflect alternate interpretations of the claim).

Although the conceptual interpretation of causal claims is the primary goal of the manipulationist account, the manipulationist account of causation carries with it an important corollary for scientific practice. For those who wish not only to evaluate the content of a causal claim but moreover to test its truth, the manipulationist account can provide norms and recommendations for experimental testing. The truth or falsity of a causal claim can be empirically tested as long as the hypothetical experiment embedded in the content of the claim

can be actually realized. Actual experiments intended to test a causal claim can—and should—be modeled on the hypothetical experiment suggested in the content of the causal claim.

Let us focus on how an actual experiment must be carried out if it is to test a causal claim:

EXPERIMENTAL INSTANCE FOR TESTING THE CLAIM “ X IS A CAUSE OF Y ”: Under some set of background conditions $\mathbf{BC} = \{BC_1, BC_2, \dots, BC_n\}$ having values $\{bc_1, bc_2, \dots, bc_n\}$, hold some set $\mathbf{S} = \{S_1, S_2, \dots, S_m\}$ of variables other than X and Y fixed at values $\{s_1, s_2, \dots, s_m\}$, perform an intervention I on X , and observe the value of Y .

The above operation, however, is only a single instance of an experiment and is insufficient for answering the question “Is X a cause of Y ?” Recall that the hypothetical experiment embodied in the claim that X causes Y makes use of a contrast between two counterfactual states: the state of Y when X is manipulated in one way, and the state of Y if X had been manipulated in a different way (or not at all). But actual experiments provide us no access to such counterfactual knowledge.

The obvious way to estimate the results of counterfactual experimental instances is to test many instances of the experimental system under similar conditions and to use statistical analysis⁵ to estimate the expected response of the system under different interventions. Let us define for this purpose an *experimental series*:

⁵Statistical analysis, as I intend it here, could be as simple as calculating a mean and standard deviation from the set of measured results, or could involve the application of much more sophisticated analysis techniques.

EXPERIMENTAL SERIES FOR TESTING THE CLAIM “ X IS A CAUSE OF Y ”: A set of two or more experimental instances for testing the claim “ X is a cause of Y ” such that:

1. Every instance in the set has the same (or sufficiently similar) values for **BC** and **S**; and
2. The set can be partitioned into two or more non-empty subsets such that every instance in each subset has the same value for the intervention I on X and no two instances falling into different subsets have the same value for the intervention I on X .

Observations made of the value of Y for each of the subsets described in item 2 above can be collated and used to generate a statistical estimate of the expected value of Y under the type of intervention used in that subset of experimental instances. If there is a significant difference in the expected values of Y for different subsets, then we may conclude that X is a cause of Y . If there is not a significant difference in the expected value of Y for different subsets, the conclusion must be more tentative. If a sufficient number of instances has been tested, we can legitimately conclude only that X is not a cause of Y under the particular circumstances of the experiment (where “circumstances” includes the background conditions **BC**, the choice of **S** on which to perform secondary interventions, and the range of values of X that were effectively tested in the series). The possibility that X will manifest itself as a cause of Y under other circumstances remains open, but the likelihood of that possibility can be reduced by testing of other series with different values for **BC**, different values for **S**, and/or interventions testing differing ranges of values of X .

4. From Real Experiments to Causal Claims. We have already discussed the way in which a real experiment can approximate the hypothetical experiment embedded in a causal claim. Now I would like to turn our attention to experiments that are *not* explicitly concerned with causation or the testing of causal claims. When analyzing an experiment that was not designed for the purpose of testing causal claims, we simply seek to identify anything that could be properly described as an experimental series (on the definition given in the previous section).

Consider as an example an experiment performed by Gasparo Berti, which aimed to decide a philosophical controversy surrounding the possibility of a vacuum and test Galileo's predictions about the maximum height to which water could be raised by suction. The experiment was most likely carried out sometime in the years 1642–1643 in the company of several active participants in the scientific scene of Rome, including Raffaello Magiotti, Athanasius Kircher, and Niccolò Zucchi. A description of the experiment is found in a 1648 letter from Magiotti to Marin Mersenne. The following is an excerpt from the letter:

In regard to the history of quicksilver, you may know that the many wells of Florence, which are cleaned each year by suction with siphons, gave Sig. Galileo the opportunity to observe the height of the attraction which was always the same, about 18 Tuscan *braccia*,⁶ and that in every siphon or cylinder, no matter how wide or thin. This was the origin of his speculations on the subject in his work on the cohesion of solids.

Later, Sig. Gasparo Berti, here in Rome, made a lead siphon that stretched about 22 *braccia* from his courtyard to his room, and was filled from above in the following way. First, leaving both valves open (D below and F above), vessel AG was filled with water. [See figure 1.] Then, after closing valve D, the water of vessel AGPM was poured out

⁶The *braccio* was equivalent to slightly more than half of a meter.

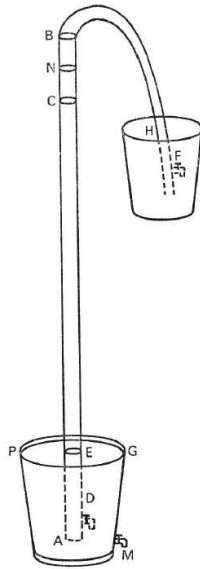


Figure 1: Diagram of Berti's experiment, included in Magiotti's letter to Mersenne

(through valve M), leaving the water inside the siphon at height AE. Later, making sure to keep vessel HF full,⁷ the water AE was allowed to flow out through valve D, which (since valve F was already open and immersed in water) pulled the water from above and filled the whole siphon BA and the vessel AG. Finally, with vessel HF full and having closed valve F, and with vessel AG full (having first closed M) and D open, the water started to descend through the siphon, emptying the entire neck BF. The water continued to fall until reaching N and did not descend further, but almost always balanced itself [at N] when the experience was replicated. And it was possible to observe this very well, since part BC of the siphon was made of glass on purpose and the whole siphon was well glued and watertight. Sig. Berti believed that he could refute Sig. Galileo with this experience, saying that the length from N to A was more than 18

⁷This was presumably done by continuous refilling.

braccia, but he should have seen that the piece of the siphon AE doesn't count, being immersed in the water of vessel AG; EN was 18 *braccia* exactly.

I should not fail to mention one thing that gave me much to think about: while the water of the siphon was falling and the neck BF was emptying, an infinite number of tiny bubbles, like those in glasses and crystals, could be seen rising through the water inside the glass BC: this, without a doubt, was some stuff that went to refill where the air was missing. I could not convince myself that it was air because there was not enough air in the water in vessel AG to refill that space (besides, the space NBF could be made much larger and it would still refill). Nor could air have entered through pores or the welding of the siphon, for if it had, it would have eventually allowed the suspended water to fall. In fact, those bubbles have always remained in my mind: I can only explain my whole sentiment about them briefly like that.⁸

Besides Magiotti's letter, there are four other sources that describe Berti's experiment: two written by eyewitnesses Zucchi and Kircher, and two other secondary sources.⁹ These other accounts all describe a similar and slightly more complex version of the experiment, which may have been a later modification. In this version, a glass globe was mounted on the siphon (see figure 2). The globe contained a bell attached to a magnetic device so that, once the purported vacuum was achieved inside the globe, the bell could be rung from outside by using another magnet.

The primary intention of the experiment, at least on Berti's part, appears to have been a desire to check (and perhaps refute) Galileo's prediction of 18 *braccia*. A secondary intention

⁸Translation mine. The manuscript of the letter is published in de Waard (1936, 178–181).

⁹de Waard (1936) contains relevant excerpts (in the original Latin) from all four sources.

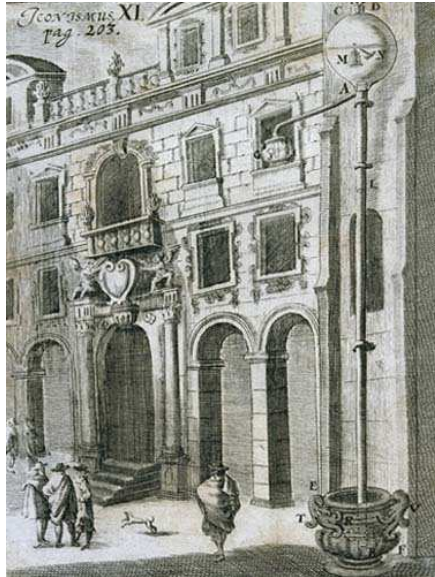


Figure 2: Engraving of a more complex version of Berti's experiment, reproduced in Schott (1664/1687, 203)

was to investigate the empty space itself: was it or was it not a vacuum? It is obvious from Magiotti's letter that this latter was a question of interest for him, and it was likely the most important question in the minds of the other participants as well; Zucchi and Kircher were both Jesuits who were convinced of the impossibility of the vacuum.

The addition of the bell in the more complex version of the experiment was suggested by Kircher and intended as an *experimentum crucis* to test the claim that the space in the globe was a vacuum. The space was found to transmit both light and magnetism, and the bell could indeed be heard when rung. These facts were enough to convince both Zucchi and Kircher, and perhaps also Berti, that the space was not a vacuum. Maignan, a friend of Berti's and a later commentator on the experiment, proposed the counter-opinion that the sound of the bell was being conducted by the bell's wooden support rather than by the space itself, and argued that the space was indeed a vacuum. It seems that Magiotti remained uncertain. Inasmuch as

the various participants walked away from the experiment with different views, the *experimentum crucis* was a failure.

Notice that the questions of interest for those performing and attending the Berti experiment were not causal questions; none of the writings explicitly mention a curiosity about the cause of the empty space, for example, nor is there any evidence of debate among the participants about what caused the elevation of the water to be 18 *braccia* rather than some other height. The questions posed and debated were, instead, factual questions and questions of interpretation about the phenomenon: How high did the water stand? Could there be any pores or imperfections in the device? Did the space transmit sound? Was the space a vacuum, or was it not?

Despite the lack of interest in causal questions on the part of those involved in the experiment, can causal conclusions can be drawn anyway? A first step toward deciding this question is to itemize the procedure described in the excerpt from Magiotti's letter and classify each step as an intervention component (I) or an observation component (O):

1. (I) Construct and set up the pipe and vessels in the configuration given in figure 1.
Ensure that valve M is closed.
2. (I) Open valves D and F.
3. (I) Fill vessel AG with water.
4. (I) Open valve M.
5. (O) Observe that vessel AG empties. Water inside the siphon remains at height AE.
6. (I) Fill vessel HF with water.
7. (I) Open valve D and continue supplying HF with water.
8. (O) Observe that the water flows out through valve D and also flows from above to fill siphon.

9. (I) Close valve F and valve M.
10. (O) Observe that the water begins to descend down the siphon, emptying neck BF and falling until it reaches N.

Assuming a similar set-up for the more complex version of the experiment,¹⁰ we might simply modify the first step and add several steps to the end of the procedure:

- 1*. (I) Construct pipe mounted with glass globe and internal magnet-bell apparatus.
Arrange it and vessels in the configuration given in figure 2.
∴
11. (O) Observe that light passes through the sphere.
12. (I) Move magnet around the exterior of the glass globe.
13. (O) Observe that the interior magnet moves in response to the exterior magnet's movement.
14. (O) Observe that sound can be heard from the bell inside the glass sphere.

It is interesting to notice that many—not just one—of the steps listed in the above procedures are interventions on the experimental system. Most of them serve only as steps toward the set-up of the apparatus. However, each can, in principle, be considered as an intervention in an experimental instance for testing a variety of causal claims; the variable X will be the thing intervened upon (for example, the intervention in step 4 is an intervention on whether or not valve M is open), the variable Y can be any observation that follows (for example, the observation in step 5 that vessel AG empties), and all other observations and

¹⁰Other accounts of the experiment describe a different procedure for filling the apparatus with water, but the difference in procedure is inconsequential for the analysis I offer below.

interventions involved in the experiment are considered either as observed background conditions in **BC** or auxiliary interventions in **S**.

The question of whether or not the experiment affords causal inference amounts to the question of whether or not the various experimental instances that make up the experiment are part of an identifiable experimental *series*. Consider, for example, an experimental instance centered around the intervention in step 4 above. The variable *X* might represent the state of valve **M** (open or closed) and the variable *Y* might represent the state of the vessel **AG** (which can be empty or full, but is observed as empty in step 5). The set-up established in steps 1–3 and other background conditions surrounding the experiment could all be represented by the set **BC**. Now, if we can identify at least one other experimental instance with the exact same values for **BC** but a different intervention on valve **M**, we will have identified an experimental series for testing the claim that the state of valve **M** is a cause of the vessel **AG** emptying. Berti's experiment does in fact provide such an experimental instance. Assuming that there is some time lapse between the execution of steps 3 and 4, we can consider as a second experimental instance the time period after steps 1–3 have been performed but before valve **M** has been opened. In this time period, vessel **AG** is observed to be full. Since there is a difference in the state of vessel **AG** between the experimental instance in which **M** is opened and the experimental instance in which **M** is not opened, we can conclude that the state of valve **M** is a cause of the state of vessel **AG**.

The observation-intervention pair considered in the example experimental series just given (i.e., a valve being opened and a vessel emptying) are such an ordinary matter of course that we do not tend to think of it as the basis for a causal conclusion that can be drawn from the experiment. That water only empties from a vessel that has some open outlet is a mundane fact that each person experiences so many times in life that it becomes an implicit piece of

causal knowledge. Still, inasmuch as the experiment establishes a contrast between performing and not performing an intervention (or alternatively, performing one type of intervention vs. performing a different type of intervention) and the corresponding difference in the observations made in each case, the experiment also affords the conclusion that one variable (the variable intervened upon) causes another (the variable observed to covary with the variable intervened upon).

But are there more substantial causal questions that could have been answered by the experiment in question? The interventions performed in the more complex version of the experiment, if compared to a relevant contrast case, could be interpreted as tests of causal questions. For example, when it is observed in step 11 that light passes through the spherical glass vessel, the implicit contrast case is whether or not light passes through the spherical glass vessel when it was originally filled with ordinary air. Presumably there were no noticeable differences between the appearance of images viewed through the vessel in the two cases. Likewise, we might compare the intervention in step 12 when it is performed in the context of the experimental set-up and when it is performed in a contrasting context (for example, with a column of water filling the siphon up to mark N, but not brought about through suction, so that the spherical glass vessel is filled with ordinary air).

The participants in the experiment were not, however, thinking in terms of these contrasting experimental instances. Even if they had been, they would have been unable to agree on a causal conclusion because they were unable to agree about what the interventions in the experiment had achieved. It is clear in Berti's experiment what the intervention is (or rather, what the sequence of interventions is: steps 1–4, 6–7, 9, 12) but what those interventions achieve was precisely the subject of debate. Some of the participants—the vacuists—thought that those interventions achieved a vacuum in the spherical vessel, while

others—the plenists—thought that the vessel was still filled with some sort of attenuated matter. If they had been able to agree, for example, that there was a vacuum in the vessel, then they might have been able to agree that ordinary air (as opposed to vacuum) was not a cause of the transmission of light or magnetism. In addition, they would have been able to reach a conclusion about the effect of the vacuum on the transmission of sound by noting any difference in the volume of the bell's ring in each case.

But there was no such agreement. Instead, some of the participants were already certain, prior to the experiment, that a vacuum could not transmit light or sound or magnetic phenomena. They took themselves to be certain of the causal relationships, and they attempted to test the presence or absence of the vacuum by the presence or absence of its purported effects. An experiment which could have been understood to test various causal claims instead used a prior confidence in those causal claims to test whether or not the cause factor was present. Even so, the actual theoretical use to which the experiment was originally put does not prevent anyone who is later informed of the details of the experiment from drawing causal conclusions.

5. Conclusion. Many experiments that are not designed for the purpose of causal inference will still afford causal inferences. The requirements I have placed on an experimental series for testing a causal claim will be found quite commonly in “ordinary” scientific experiments. We can see that this is true especially when we consider that, in cases where there is a time lapse between the set-up of the experiment and the intervention on the purported cause variable (if the time latency of the observed result is small in comparison to the time lapse), a comparison of observations made before and after the intervention is performed will usually correspond to an experimental series for testing if the variable intervened on is a cause of the

subsequent observation.

Interestingly, the fact that many “ordinary” experiments will afford causal inference means that any experimental science has a plentiful source of causal content. I see this unacknowledged point as significant to debates about whether or not there is causal content in fundamental physics.¹¹ In acknowledging the epistemic dependence of fundamental physics on experiment, we must also acknowledge at least a potential for causal content.

¹¹For a set of papers in this debate, see the volume edited by Price and Corry (2007).

References

- de Waard, Cornélius. 1936. *L'Expérience Barométrique: Ses Antécédents et Ses Explications*. Thouars: Imprimerie Nouvelle.
- Price, Huw and Richard Corry, eds. 2007. *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*. Oxford: Clarendon Press.
- Schott, Gaspar. 1664/1687. *Technica Curiosa, Sive Mirabilia Artis*. Herbipol.: Jobus Hertz.
- Spirtes, Peter, Clark Glymour, and Richard Scheines. 2000. *Causation, Prediction, and Search*. 2nd ed. Cambridge, MA: The MIT Press.
- Woodward, James. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.
- Zhang, Jiji and Peter Spirtes. 2008. "Detection of Unfaithfulness and Robust Causal Inference." *Minds and Machines* 18 (2): 239–271.