Review of Woodward, *Making Things Happen*

Michael Strevens

Draft of December 2006

A causal account of scientific explanation requires a dual foundation: a theory of causality, and a theory of what causal facts are put together in what way to create explanations of events, regularities, laws, and so on.

Jim Woodward's *Making Things Happen* has much to offer on both counts. It provides a theory of causation, a theory of event explanation, an account of what Woodward calls invariant causal generalizations, and a welcome discussion of the causal interpretation of structural models (such as linear causal models), as well as extended discussion of many facets of the literature on scientific explanation.

The theories of causation and event explanation constitute the core of the book. If you accept what Woodward says about these topics, you will very likely agree with what he has to say about invariant generalizations and structural models. This short discussion will therefore focus on the core.

Woodward is a manipulationist both about the nature of causation and about explanation. Despite the terminological overlap, causal manipulationism and explanatory manipulationism turn out to be distinct and quite independent views. (Causal manipulationism in turn has two independent parts, one pertaining to type level, the other to singular, causation.) Section 1 of this essay sketches Woodward's causal manipulationism, putting aside the technical details; section 2 discusses Woodward's manipulationist account of event explanation; section 3 returns to and critiques the theory of type level causation.

1. Correlation, Manipulation, Intervention, Causation

What does the claim that coffee-drinking causes cancer say over and above the claim that cancer is correlated with coffee-drinking? According to Woodward, that you can affect your chances of cancer by changing your coffee consumption. Thus, if there is correlation but no causation—if, for example, the statistical connection is due to some hidden factor, say stress, that causes both coffee-drinking and cancer—then you cannot increase your chances of getting cancer by drinking coffee. What causal information adds to information about correlation is, in a word, information about manipulability.

This characterization must, however, be carefully circumscribed. In the common cause case, if you adopt the indirect strategy of upping your coffee intake by subjecting yourself to greater stress, then you *will* increase your chances of getting cancer. For the purposes of characterizing causality, then, only certain techniques for drinking more coffee are allowed. These Woodward and other manipulationists call *interventions*.

Without too much distortion, you can think of an intervention in Woodward's sense as the result of the hand of God descending and directly tweaking the relevant factor—filling you with coffee without your even noticing, for example. Or you can think of it as a change that results from a kind of Lewis-style "small miracle". Either way, your coffee intake (or whatever it is) is altered with the fewest possible side effects, ideally none. It is only when this sort of change allows the manipulation of the putative effect that you have genuine causation.

What is causal manipulationism? It is not merely the position that facts about causal relations entail and are entailed by facts about manipulability. A close relation between manipulation and causation is compatible with many different accounts of causality. What is characteristic of manipulationism is the doctrine that the facts about causation *metaphysically depend* in part on the facts about what can be manipulated by what.

Manipulability theories of causation are not new; Woodward's version is distinctive (if not unique) in two ways. First, it draws heavily on the interdisciplinary literature on Bayesian causal networks, a body of work until recently

focused on causal inference—both from causes to effects and effects to causes (Pearl 2000; Spirtes et al. 2000)—rather than the metaphysics of causation. This gives Woodwardian manipulationism considerable technical firepower (used in this book with a very light touch). Second, Woodward is more concerned than earlier manipulationists (Menzies and Price 1993, in particular) to fashion an account on which what counts as an intervention, hence what counts as causation, is an entirely observer-independent matter. Woodward's criterion for being an intervention has nothing to do with human causal capabilities: there are technically correct interventions that no human will ever carry out, perhaps even that are physically impossible to carry out (p. 130).

On Woodward's view—typical in philosophy of science—causation is in the first instance a relation between types, and only in a derived sense a relation between particulars. The preferred type is what Woodward calls a *variable*, meaning rather what is represented by a variable in a causal network—namely, a determinable such as electrical charge, fecundity, or income. The values of a variable are, then, determinates of the determinable. (Perhaps it is better to think of a variable as a function from a determinable's determinates onto a set of numbers, like a random variable in probability theory.) Variables need not be continuous or even many-valued: a binary variable takes only two values, and may be thought of as representing the occurrence or non-occurrence of a given event type (or the obtaining or not obtaining of a given type of state of affairs).

One variable causes another, very roughly, if intervening on the value of the one sometimes changes the value of the other (for refinements, see section 3). The type level relations between different variables can be represented in a directed graph, in the mathematical sense of *graph*; an example is shown in figure 1. The manipulationist's project may be understood, then, as an attempt to characterize the nature of the facts represented by such a graph.

Technical literature on Bayesian networks may limit or augment a causal graph in various ways, thereby subtracting from or adding to the power of the graph to represent different aspects of the causal world. An important, but not compulsory, limitation is the requirement that the graph be acyclic, that is, that it be impossible to trace a path out from a variable and back to itself following

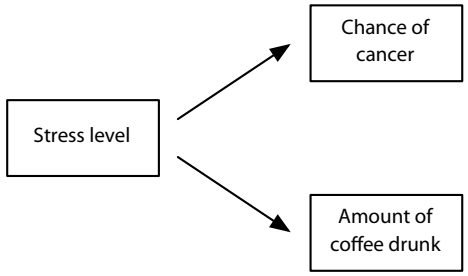

Figure 1: Simple directed graph

the causal arrows in one direction only. An important augmentation is the imposition of a probability distribution over the graph, allowing the graph to represent not just causal connections but the strength and interaction of those connections. This augmentation is itself usually constrained, in particular by the Markov condition, which requires that the structure of the graph be reflected in certain independences in the probability distribution. Woodward for the most part focuses on the simplest case, that of the interpretation and explanatory use of acyclic, deterministic graphs. I will raise the question of generalizability in section 3.

Causal graphs figure in three ways in Woodward's accounts of causation and event explanation. First, his theory of type level causation is a theory of the nature of the causal connections between variables, that is, of the nature of the facts that are represented by the arrows in a causal graph. Second, his theory of singular causation consists of a method for extracting the facts about the "actual causes" of a singular event from an appropriate causal graph. Third, his theory of event explanation is an account of how to use the actual causes of an event so extracted, along with other information also drawn from the graph, to explain the event. I will discuss these elements in reverse order, the latter two in section 2 and the first in section 3.

2. The Theory of Event Explanation

Explanatory Manipulationism To explain a phenomenon, on Woodward's view, is to show how to manipulate the occurrence of that phenomenon. Explanation is not, as you might have thought, a higher and less practical endeavor than prediction; on the contrary, the explainer in a certain sense takes a more active and less spectatorial stance to the world than the predictor.

To show how to manipulate a phenomenon is to show how that phenomenon would change as various interventions are carried out on its causes. Thus it is to exhibit a "systematic pattern of counterfactual dependence" of the explanandum—the event to be explained—on various causal factors (p. 191). For this reason, Woodward calls his account a counterfactual theory of explanation, but keep in mind that only a certain species of counterfactuals have explanatory power, namely, those pertaining to the effects of interventions.

Although Woodward's formal characterization of explanation applies only to events (p. 203), he also has something to say about the explanation of generalizations. However, the reader is for the most part left to extract the rules of generalization explanation from examples. In the simple cases described in detail on pp. 187–194, this is easy enough; it is harder to see how causal networks are to be deployed in the more complicated cases mentioned in passing in the book, such as the use of kinetic theory to explain the ideal gas law. For this reason, my focus will be Woodward's account of the explanation of events.

Singular Causation The core of Woodward's account of singular event explanation is the account of singular event causation, which has much in common with accounts of singular causation offered by other Bayes nets aficionados (Halpern and Pearl 2005). Let me explain its workings by contrasting it with a simple counterfactual account of causation, on which an event *c* is a cause of an event *e* just in case if *c* had not occurred, *e* would not have occurred. To highlight the strengths of Woodward's approach, consider the sort of preemption scenario known to create difficulties for the simple counterfactual account.

The classic preemption scenario is the case of the wavering assassin. In a Middle Eastern country near you, a rebel is deputized to activate the detonator

when the general's motorcade passes by the bomb. Because the rebel has qualms, another rebel of firmer resolve is posted further down the road with a rocket launcher. If the wavering rebel fails to detonate the bomb, the resolute rebel will fire the rocket. All causation is deterministic: if either the bomb is set off or the rocket is launched, the general is sure to die, and if the bomb is not set off, the resolute rebel is sure to launch the rocket.

As it happens, the wavering rebel goes ahead and detonates the bomb, killing the general. According to the simple counterfactual account, however, the rebel's hitting the button on the detonator is not a cause of the general's death, because it fails the counterfactual test: had the rebel not pressed the button, the general would have died anyway, by rocket.

Woodward's treatment of such a case begins with a causal network—a directed graph—representing the causal connections involved. My version of the network (figure 2a) uses only binary variables, representing the occurrence or non-occurrence of various events, such as the wavering rebel's pushing the button. The arrows between the boxes represent the causal consequences of the variables: an arrow with a plus sign represents the effect of an event's occurrence, an arrow with a minus sign the effect of its non-occurrence. The network is a type level model: it represents the causal structure of any scenario answering the description above, with the values of the variables representing, for any instantiation of the scenario, the way that things pan out.

Figure 2b shows what happens in the actual case under consideration, with a slashed variable representing the non-occurrence of the corresponding event and an unslashed variable representing its occurrence. As shown, the button is pushed and the bomb explodes; the trigger is not pulled and the rocket is not launched. Dashed arrows represent causal connections that are not actualized.

The application of the simple counterfactual test for singular causation can for expository purposes be approximated using the causal network in the following way. For a putative cause such as the pressing of the button, switch the corresponding variable so that it represents the event's not occurring. Technically, this switching constitutes an intervention in Woodward's sense, since it controls the value of the switched variable directly and has no side effects.

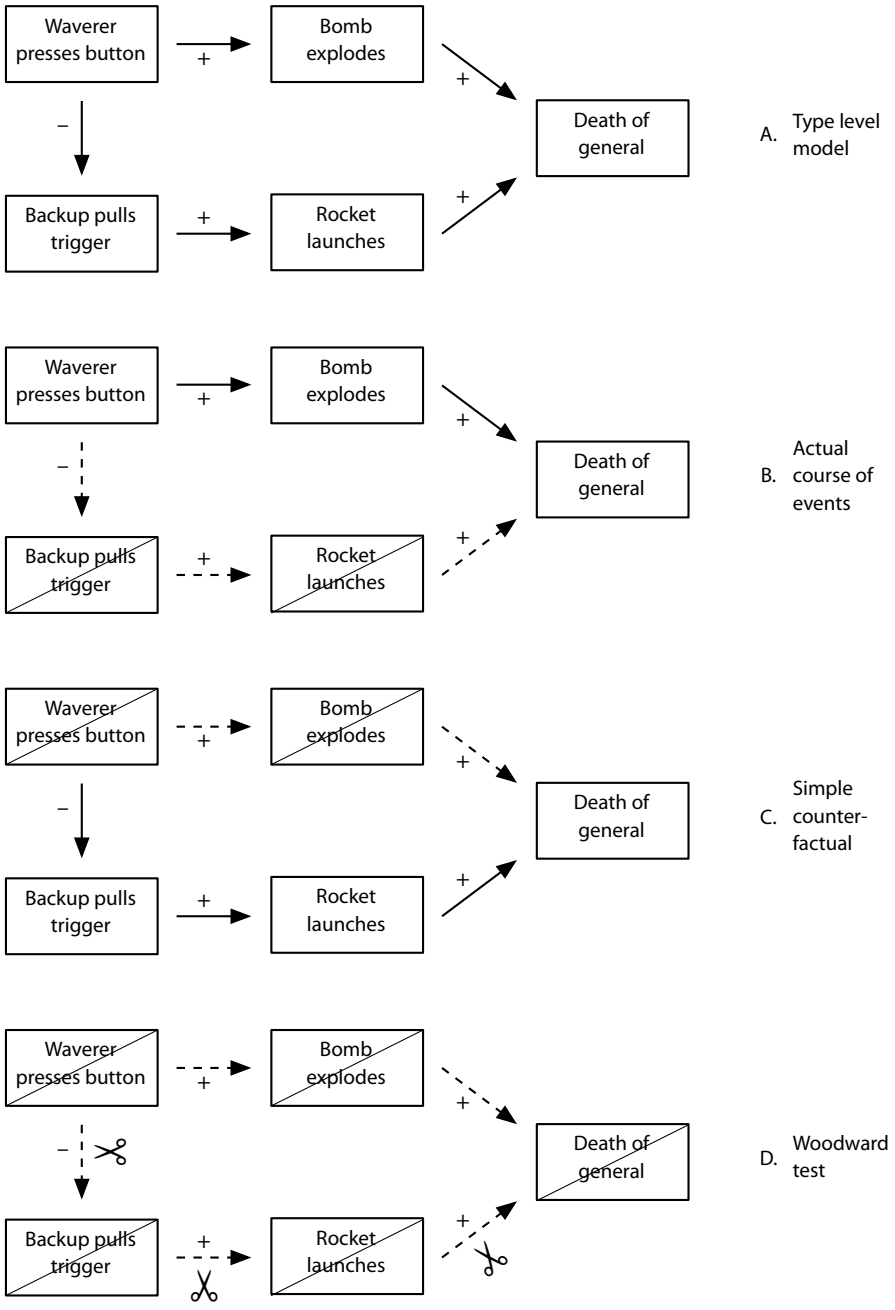

Figure 2: Event explanation using a causal network: two approaches

Hold all other initial conditions—variables with no incoming arrows—fixed at their actual values (a stipulation with no effect in the simple networks shown in figure 2, since there is only one initial condition). Allow the effects of the switching to propagate through the network. Your putative cause is a genuine cause just in case the explanandum event does not, according to the network, then occur. As shown in figure 2c, conducting the counterfactual test on the button-pressing does not change the value of the *death* variable, and so the button-pressing is falsely declared not to be a cause of death.

Woodward's test is like the simple counterfactual test, with one crucial amendment. You do not simply test an event for causehood; you test an event together with a putative causal path. As in the counterfactual test, you switch the event "off" using an intervention, but in addition, you hold all variables not on the chosen path to their actual values. An event is a cause of the explanandum event by way of the chosen path just in case the explanandum occurs when the putative cause is switched on, but not when it is switched off.

There is a simple algorithm for holding the variables off the chosen path to their actual values without creating inconsistencies: you isolate them from all incoming causal influences by severing incoming causal arrows in the network. You can then allow the effect of your switching off the tested event to propagate through the network without worrying that variables that are supposed to be held constant will be affected. The process is shown, for the button-pressing and the path from button-pressing to death via the bomb, in figure 2d. Scissors indicate the severed arrows (in this case only the first really needs to be severed). Because the behavior of the backup rebel is not on the tested path, it is held fixed at its actual value, so that the rocket is not launched when the button-pressing is switched off. Thus the general lives, and the button-pressing is declared a cause.

What the simple counterfactual account does wrong, you will observe, is to allow the activation of the backup cause—the resolute rebel—in the causal network used to evaluate the relevance of the actual cause. Woodward's approach succeeds where the counterfactual approach fails because it does not allow the backup cause to operate.

As well as providing an attractive solution to preemption problems, Woodward's procedure has the virtue of picking out a particular set of causal generalizations, alongside events, as the explainers of any given event, namely, the generalizations in virtue of which the implicated causal pathway exists. For a manipulationist, the generalizations are in many ways the most important part of an event explanation, since they are the repository of information about manipulability.

Observe that on Woodward's account of singular causation, much of the work in deciding what is causally relevant and what is not is done by the facts about type level causation, rather than by the procedure outlined above, since it is the type level facts that individuate causal pathways, and matters of causation turn on questions of pathway individuation. In the assassination case, for example, the wavering assassin exerts various kinds of causal influence on the backup assassin, and vice-versa—perhaps they can see one another; certainly there is a (very small) gravitational influence between them. If, in virtue of this mutual influence, the backup assassin were to be counted as part of the causal pathway leading from the wavering assassin's button-pressing to the general's death, then the backup's actions would not be held fixed in the evaluation of the button-pressing's causal status, and so the Woodward test would fail in the same way as the counterfactual test.

It is crucial for the integrity of the Woodward test, then, that the gravitational influence and so on be determined to be causally irrelevant before the test begins. It is the high level, type level relations represented by the network that fix such facts, by determining that the gravitational influence plays no *relevant* role in the causal connections between button-pressing and death, despite its undeniable low level causal involvement in the connections. As a result of this determination, the gravitational influence is simply left out of the causal network. This raises the question whether Woodward's metaphysics of type level causal relations, to be discussed in section 3, is subtle enough to make these distinctions of relevance.

Trouble with Preemption Even if the type level causal facts do everything that can reasonably be asked of them, Woodward's account of singular causa-

tion appears to fail in cases where the backup causal pathway is also a part of the actual causal pathway, thus the backup cause is also an actual cause. In such a case, the values of the variables on the backup pathway are not held fixed; unfixed, however, they are free to play the backup role, destroying the counterfactual dependence of the effect on its other actual causes.

Suppose, for example, that Radio Rebellion broadcasts the code word calling for the wavering rebel to detonate the bomb. The resolute rebel, hearing the code word and knowing that the rebel is wavering, shows herself. The wavering rebel, seeing the resolute rebel in backup position and reasoning that the general will die whether he detonates the bomb or not, overcomes his moral discomfort and pushes the button. Assume that the resolute rebel would not have showed herself if the code word had not been broadcast, and that the wavering rebel would not have pushed the button if the resolute rebel had not been in position to launch the rocket. Is the broadcasting of the code a cause of the general's death?

As I have told the story, the elements of the backup pathway—rebel plus rocket launcher—are also a part of the actual causal pathway leading from the broadcast to death, since it is partly in virtue of its actual effect on these elements that the broadcast is able to cause the wavering rebel to press the button and so to cause the general's death. But since these elements lie on the actual pathway, the variables that characterize their state in a causal network cannot be held fixed. When the effects of switching off the ordering propagate through the system, the backup is activated and the rocket launched, as in the simple counterfactual test (figure 2c). The general dies anyway and the broadcast is therefore declared not to be a cause of the death.

This is, because the causal relations involved are difficult to parse into a causal network, not the perfect counterexample to Woodward's account of singular causation, but it is easy to use levers and buttons, or if you prefer logic gates, to construct a mechanism in which the very same sub-mechanism is both triggered by and necessary for an actual cause to do its work and in the absence of that actual cause (and only in its absence) sufficient to do the same work.

Explanatory Depth Some event explanations, according to Woodward's explanatory manipulationism, are better than others, because they convey more manipulative know-how. (Woodward rather unconventionally identifies the amount of manipulative know-how conveyed with the depth of the explanation.) Suppose that the event to be explained is some particular object's having a certain property. The more the explanatory causal network tells you about ways in which you can manipulate the object so that it has the same or similar properties, the better (the "deeper") the explanation.

Thus, although the information about the actual causes of an event *e*, along with information about the causal pathways in virtue of which they did their actual causing, derived using the procedure described above, constitutes the core of an explanation of *e*, you have an explanatory duty to augment this core. The best explanation of *e* will contain, not only information as to how the occurrence of *e* was in fact caused, but also information as to how *e* might have been but was not caused, and how other similar but distinct events might also be caused. It is this doctrine, more than any other, that distinguishes explanatory manipulationism from rival causal approaches to explanation.

Does this let backup causes such as the resolute assassin's trigger finger, so painstakingly stripped of causal status by the manipulation test described above, in through the explanatory back door? Woodward is reluctant to say (pp. 219–220). Yet surely information about a backup cause constitutes a way that the effect might have been, but was not brought about, hence constitutes relevant manipulative know-how? On Woodward's proposal, then, it seems that backup causes, though quite deliberately declared not to be actual causes, nevertheless help to explain the event that they did not cause.

If this is not to count as a philosophical own goal, Woodward needs a good argument for including in the explanation of an event information about these non-actual causes. Perhaps his best hope is the observation that the property of robustness counts for something in causal explanation. A process producing an event *e* is robust if a wide range of initial conditions besides the actual initial conditions would have led to *e*; observing, in the course of explaining *e*, that the process producing *e* was robust adds to our understanding of why *e* occurred—or so most writers (including this one) would agree.

To get from the explanatory importance of robustness to manipulationism requires two further steps. First, it must be shown that information about robustness goes beyond information about how *e* was actually caused. This is far from clear: in the paradigms of explanatory robustness—processes that robustly produce an event of a certain type *e*, and in which the robustness clearly helps to explain *e*, for example, statistical mechanical processes (Railton 1981)—the features of a causal network responsible for the robustness are arguably identical to the features responsible for producing any particular instance of *e*; they are the very same type level causal relations. Thus a delineation of the actual causal process leading to *e* is also a delineation of the explanatorily valuable facts about robustness. No additional causal information is needed.

Second, rival approaches to causal explanation that claim, like explanatory manipulationism, to make sense of the need for information over and above information about actual causes—for example, Jackson and Pettit (1992)—must be debunked. Woodward does very little to establish either step. He claims that information about robustness is not contained in information about actual causation (p. 232), but does not consider any alternative view. And there is no systematic comparison of explanatory manipulationism's treatment of robustness with that of other approaches to causal explanation.

Variable Relativity On Woodward's account the causal relationships between variables in a causal network may change as variables are added, with arrows both disappearing and appearing.

Assume, for example, that the amount of expensive bottled water you drink and your chances of succumbing to heart disease are correlated, because they share a common cause, say, the consumption of salty food. Consider a causal network containing variables representing water-drinking and heart disease but not salt consumption. Because Woodward's definition of an intervention is implicitly relativized to the variables in a network, increasing the amount of bottled water you drink by increasing your consumption of salty foods will count as an intervention relative to the salt-free network (due to the invisibility, within the network, of the "side effects" of the salty strategy for drinking more).

Thus, because "intervening" on your water consumption in this particular way will increase your chances of getting heart disease, water consumption will count as a cause of heart disease, or in other words, the network will include an arrow from water consumption to heart disease.

To make sense of this phenomenon, Woodward chooses to relativize the notion of causation to a set of variables. He would say, in the case above, that drinking bottled water causes heart disease relative to a set of variables that does not include consumption of salty food, but does not cause heart disease relative to a set that does include salty food. Never mind for now whether this is good causal metaphysics; does it yield a good account of explanation? Only if there is some strategy for choosing a set of variables relative to which to answer any particular explanatory request. The obvious strategy is to choose all the variables.

Woodward cannot make this move. Because of the way that he defines causation, a causal network must contain "shortest causal links", what Woodward calls relations of direct causation. But if your causal process is continuous, and you have variables representing every aspect of the process, then you can have no shortest causal link. Between any two variables lies a further, intermediate variable. (There is an additional problem concerning variables that represent the world at different levels of detail.) Thus a Woodwardian causal network cannot contain "all the variables". Indeed, Woodward's entire causal apparatus, and his notion of direct causation in particular, is founded on the supposition that causal networks represent less of causal reality than is actually out there.

What Woodward urgently needs, then, is a criterion capable of picking out sets of variables that are, on the one hand, sparse enough to enable the definition of a relation of direct causation, but on the other hand, complete enough not to give rise to spurious explanation, such as the explanation of heart disease by bottled water in the case above.

As far as I can see, there is no such criterion on offer in *Making Things Happen*: "the matter deserves more detailed treatment than I can give it here" (p. 80). Woodward does write that causal and explanatory claims should be evaluated only against variables representing "serious possibilities", but this po-

sition is tailored for an orthogonal worry about the causal status of omissions, and does not seem well suited to solving the more pressing problem. (In any case, since the facts about what causes what themselves depend on the variables chosen, it is rather unclear by what token a variable becomes a serious possibility—it cannot have anything to do with causal relevance.)

3. Metaphysical Foundations

It is the type level causal relations, on Woodward's view, that provide the metaphysical basis of causal explanation, by determining the facts about singular explanation and causal relevance. Woodward promises to give an account of "the content or meaning" of statements of type level relations, in particular the paradigm *x causes y* (where *x* and *y* are variables; p. 38). There are reasons to doubt that what he provides fits this description, at least if the content of such claims is supposed to fix their truth conditions.

All causal relations between variables are built from the relation of direct causation, which is the relation represented by an arrow in a causal network. A variable *x* is a direct cause of another variable *y*, relative to a variable set *V*, just in case there is an intervention on *x* that will change the value of *y* (or the probability distribution over the values of *y*) when all variables in *V* except *x* and *y* are held fixed (p. 55). When this definition is satisfied, no other variable in *V* acts as a causal intermediary between *x* and *y*. (More exactly, there will be at least one path with no intermediaries, that is, one arrow going straight from *x* to *y*.) Thus if the world's causal processes are continuous—in which case all causal relations are mediated—the notion of direct causation is only applicable relative to a *V* that represents a suitably sparse subset of the totality of causal facts.

It is hard to know what to make, then, of Woodward's view that direct causation is the fundamental causal notion. Is the claim that causal processes are continuous to be understood as a second-order property of the facts about direct causation?

The relativization of the relation of direct causation is also a puzzle. Our

notion of causation is surely not relativized to a variable set. Thus, direct causation would have to be a term of art, presumably intended to help define a non-relativized notion of causation that captures the "content or meaning" of our causal claims. But no such non-relativized notion appears in *Making Things Happen*.

Finally, consider the fact that Woodward, having defined direct causation using the notion of intervention, self-consciously defines intervention using the notion of causation. If the causal manipulation of a variable x is to count as an intervention on x relative to a set of variables V, and is thus to be useful in determining what is caused by x relative to V, several conditions must hold, all mentioning causation (p. 98). For example, the manipulation must have no effect on the variables in V that does not go by way of x, which is to say that every causal pathway from the manipulating apparatus to the variables in V must go by way of x.

As you can immediately see, it is impossible to determine whether a manipulation is an intervention on x relative to V without knowing about the causal pathways that connect the members of V, which is to say, without consulting a causal network for V. But to have a causal network for V is to have settled all questions about direct causation between the members of V, and to answer these questions, you must know what counts as an intervention on each of the members of V.

Woodward denies that his definitions are literally circular, in the sense that the fact as to whether x causes y depends on what is caused by x, but it seems that in some cases, there is exactly such a dependence. Suppose that I have a set V of three variables, x, y, and z, and I wish to know whether x is a direct cause of y. I need an intervention on x. This intervention (itself represented by a variable i) must not be causally linked to any variable that is in turn a cause of y, unless the link goes by way of x. In particular, either it must not be a direct cause of z, or z must not be a direct cause of y, in each case relative to V. To rule out the first possibility, I have to find an intervention on the intervention variable i. Infinite regress, you are thinking. Actually, worse: even if the gods were to hand me such a variable, I am now testing for causation relative to a variable set that includes not only the original three variables but

i as well. What I needed, however, were the facts about causation relative to just the three variables. Dead end. So perhaps I can show that *z* is not a direct cause of *y*? To do this I need, by parity of reasoning, to settle the question of whether *x* directly causes *y*. Circularity.

I have put things in epistemic terms, but Woodward is supposedly providing a set of truth conditions for causal claims. It is hard to see how, on his view, such claims could possibly be grounded.

One way out of this bind is to remove the circularity in the definitions. The conventional way to do so is to define intervention in terms of type level causation, as Woodward does, but then to define type level causation in some way that makes no reference to intervention, or to take it as a primitive. This is, of course, to renounce a manipulationist metaphysics of type level causation. (A non-manipulationist conception of type level causation is standard in the Bayesian networks literature. This approach in no way precludes the use of interventions, Woodward-style, to define singular causation (Pearl 2000).) The alternative is to define intervention non-causally, perhaps by invoking human paradigms of manipulation. This is, obviously, not to Woodward's taste.

Woodward would be much better off, I suggest, presenting his account not as a metaphysics or a definition of type level causation, but as describing something like a Putnam-style stereotype of our causal concepts, that is, as a theory of a kind of "content" that does not fix truth conditions. The view might go as follows. There is a fact of the matter about type level causal relations that has nothing to do with interventions. (Choose your favorite causal metaphysics.) We humans have concepts and terms that succeed in referring to causal relations, but the internal aspect of these concepts, the part that is "in the head", falls far short of specifying the true nature of causality. It is this internal aspect that is captured by the Woodwardian definitions. You might say, with profuse apologies to Kant, that manipulationism as a metaphysical doctrine is true of the phenomenal world, but false of the noumenal world.

However, even this view cannot, I think, be entirely correct. There are central aspects of our conception of causation that are not captured by Woodward's definitions. In particular, we hold that lower level causal relations constrain higher level causal relations: the fact that smoking causes cancer is not

compatible with just any arrangement of low level causal relations (Strevens forthcoming). Because Woodward's definitions relate facts about high level causation to other facts (sometimes the same facts) about causation at an equally high level they have, as far as I can see, no implications for interlevel dependence—and so they do not give a complete account of the properties attributed to causal relations by our conceptual apparatus.

One more comment. Woodward is committed, on either the noumenal or the phenomenal interpretation, to the adequacy of the directed graph as a representation of all causal relations (though not, of course, to the existence of a single graph capturing all causal relations).

This commitment is dubious for several reasons. First, directed graphs are clumsy representers, or worse, of many kinds of causal processes—try to draw a causal graph for my signing my name, or the initial break in a game of pool. Second, by their very nature, they divide processes that we believe to be continuous into discrete steps, an imposition that is sometimes fruitful, but sometimes not. Third, they are quite unsuited to representing the dependency between high level causal relations and the low level relations on which the high level relations supervene.

Perhaps every causal fact can, in principle, be represented by a causal graph, but even so I suspect that Woodward, in his making "direct causation" basic and in relativizing causation to a set of variables, has taken certain aspects of this particular representational tool too literally, projecting them with little warrant onto causal reality.

4. Assessment

I have examined three aspects of *Making Things Happen*: its account of type level causal relations, its account of singular causation, and its account of causal event explanation. Each is a view that can rightly be labeled *manipulationist* (in the case of the accounts of type level and singular causation, because of the uses they make of the notion of intervention). But as remarked above, each is independent of the others: you can consistently adopt any

combination of the three that you like. I will evaluate them separately.

The manipulationist account of type level causal relations is, I think, a strategic error, introducing as it does complex philosophical quandaries which would perhaps take their own dedicated book-length study to resolve. In any case, Woodward's allegiance to type level manipulationism appears to be based on a misapprehension. He considers it perplexing that philosophers have on the whole rejected manipulationism whereas many scientists and statisticians consider the connection between causation and manipulation to be central to the scientific method. Siding with science, Woodward chooses manipulationism (pp. 25–28). But an examination of the various quotations and other considerations that Woodward assembles to demonstrate the scientific significance of the causation/manipulation connection shows that what is important is that facts about causation should entail facts about manipulation and (perhaps) vice-versa. The recognition of such a connection is entirely compatible with the rejection of metaphysical manipulationism, which is a matter not just of entailment but also of ontological dependence. (I should note that some of Woodward's sources make the same conflation.)

My verdict: Woodward's type level metaphysics should be abandoned, along with the relativization of causation. Everything else in the book, in particular the use of graphs as causal models, the manipulationist definition of singular causation, and the manipulationist theory of explanation, works perfectly well with a non-manipulationist conception of type level causation.

The definition of singular causation is not perfect, as the counterexample in section 2 shows, but it is powerful, sophisticated, and usually right—one of the best in the literature (though naturally, I prefer the definition presented in Strevens (in press)). My main concern is that it offloads so much of the work in determining what is relevant to the occurrence of an event to the high level, type level causal relations. Given a question as to how an event was caused, the type relations determine the possible causal pathways; to determine the actual causes, you need only look to see which of the pathways was in fact instantiated. If there are indeed high level causal relations covering every case of singular causation, well and good, but there currently exists, I think, no account of such relations powerful enough for Woodward's needs.

The manipulationist account of explanation is certainly provocative and interesting, most of all in its claim that the best event explanations not only present the actual causes and causal pathway responsible for the explanandum event's occurrence, but also ways that it might have been but was not caused, and even ways that other similar events might have been but were not caused. The question that non-manipulationists will put to Woodward is as follows: why, once you have given the complete story as to how an event was actually caused, do you improve an explanation by adding further information about non-actual causation?

It is surprisingly difficult to find an answer in *Making Things Happen*. First, as remarked in section 2, Woodward declines to join the debate about robustness.

Second, Woodward devotes little effort to distinguishing the undeniable practical utility of adding supplementary information about manipulability to an explanatory model from the question whether such an addition increases the model's explanatory power.

Third, although the supplementary information is said to increase the explanation's depth, Woodward does not examine any cases in which his criterion for depth diverges in its judgments from a more traditional criterion on which depth is a matter of describing causal processes in terms as close as possible to the level of fundamental physics. Rather, his strategy is to argue that the two will not in general come apart (§5.10), thus depriving readers of a potentially decisive test case.

Finally, the question of how much, and what kind of, manipulative know-how should be added to an explanation above and beyond the information about the actual causation of the explanandum is, though addressed many times, never answered with the level of precision that, say, Hempel would have demanded.

As a result, the reader is left with a good sense of what it is like to be a manipulationist about explanation, and of the characteristic moves that an explanatory manipulationist will make in response to various objections, but this book is unlikely to convince anyone not already sympathetic to join the movement.

Perhaps the most valuable contribution of *Making Things Happen* is independent of all three of Woodward's manipulationist doctrines: it is the insight the book provides into the scientific role of the connection between causation and manipulation. How, to pose just one interesting question, does it change our understanding of generalizations in the special sciences if their empirical content concerns not just correlations but interventions? Read the chapter on structural equations for the beginnings of an answer. But do not make the mistake of thinking that its importance is reason to philosophically apotheosize manipulation as the foundation of everything—or perhaps anything.

References

Halpern, J. Y. and J. Pearl. (2005). Causes and explanations: A structural-model approach. Part 1: Causes. *British Journal for the Philosophy of Science* 56:843–887.

Jackson, F. and P. Pettit. (1992). In defense of explanatory ecumenism. *Economics and Philosophy* 8:1–21.

Menzies, P. and H. Price. (1993). Causation as a secondary quality. *British Journal for the Philosophy of Science* 44:187–203.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge.

Railton, P. (1981). Probability, explanation, and information. *Synthese* 48:233–256.

Spirtes, P., C. Glymour, and R. Scheines. (2000). *Causation, Prediction, and Search*. Second edition. MIT Press, Cambridge, MA.

Strevens, M. (Forthcoming). Why think causally? In A. Gopnik and L. Schulz (eds.), *Causal Learning: Psychology, Philosophy, Computation*. Oxford University Press, New York.

———. (In press). Mackie remixed. In J. Keim Campbell, M. O'Rourke, and D. Shier (eds.), *Causation and Explanation*, volume 4 of *Topics in Contemporary Philosophy*. MIT Press, Cambridge, MA.