Pragmatic Causation[*]

ANTONY EAGLE[†]

Abstract Two arguments due to Russell are examined, and found to show that the notion of causation as full determination doesn't mesh easily with deterministic global physics and the distinction between effective and ineffective strategies. But a local notion of causation as involving a certain kind of counterfactual dependence is, I argue, compatible with Russell's conclusions. I defend it from a resurgent form of Russell's microphysical determinism argument by deploying a pragmatic account of the nature and function of scientific theories.

1. Russell's Arguments

> The law of causality. . . like much that passes muster among philosophers, is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm. (Russell, 1913 [1963], 132)

Russell (1913 [1963]) takes the relation of causation to be a relation of determination: c causes *e* just when c determines *e* to occur. This relation is asymmetric and plausibly transitive. The fundamental law of causality is supposed to be that every event has a sufficient cause, one that is guaranteed to bring that event about and in fact did so. This intuition about the deterministic nature of causation is not a Russellian idiosyncrasy: it originates in Hume's 'constant conjunction' regularity analysis (if c and *e* are *constantly* conjoined, the appearance of c should

[*]Thanks to audiences at Princeton and at the *Causal Republicanism* conference, Centre for Time, University of Sydney. Particular thanks go to Huw Price, Chris Hitchcock, Helen Beebee, Daniel Nolan, Gil Harman, Charles Twardy, Karen Bennett, Mathias Frisch, Jason Grossman, Gill Russell, Paul Benacerraf, Mark Schroeder, Graham McDonald, Brett Sherman, and Adam Elga. Special thanks to Toby Handfield and Jeff Speaks.

[†]Philosophy Department, Princeton University. aeagle@princeton.edu

be sufficient for the appearance of *e*), and even later accounts like Suppes (1970) keep the idea that individual causes partially determine their effects, and deny that every event has a sufficient cause to avoid the supposed 'universal law'.

Russell thinks that this notion of causation as a determination relation between events doesn't appear in physics, and hence should be jettisoned from a properly scientific world view. Field (2003) identifies two arguments in Russell to this conclusion. The first rests on the claim that the equations of microphysics are bi-deterministic. If we fix the global microphysical state *s* of some system R at t, that fixes the whole trajectory of R through the space of states both before and after t.[1] If all macroscopic events are constituted by some particular microphysical states, then fixing a particular event in the system at a time will determine which microphysical trajectory the system is on, and hence which events will occur and have occurred. Then any event determines both its temporal antecedents and temporal succedents. But if 'c causes *e*' is defined in Russell's sense—namely, c determines *e*—then by this argument *e* equally well causes c. The asymmetry of determination, and hence the asymmetry of cause and effect, is lost. There seems no place in bi-deterministic physics for the causal asymmetry.

This argument has at least two flaws. Firstly, though it might be true that the causal asymmetry is not an asymmetry of determination, causation still might be defined from a relation of determination *combined with* an asymmetrical relation (where the asymmetry comes from somewhere else). Perhaps the asymmetry is that of entropy increase, plausibly set by unusually low entropy initial conditions.[2] Perhaps it is a brute temporal asymmetry.

Secondly, it may still be the case that there is a macroscopic asymmetry between causes and effects. Perhaps on a global scale, the whole state of the universe at one time determines the whole state at every other time. However, if we restrict our attention to a local particular event c, presumably there are many global states that can have this event as a constituent part.[3] Perhaps all the global trajectories which feature this event in some state have some further event *e* as a feature of a state; but perhaps not all the global trajectories which feature *e* have c as a feature. So the occurrence of c determines the occurrence of *e* in a way that *e* doesn't determine the occurrence of c. Focussing attention on a limited area might very well give us an asymmetry of determination between particular events. Moreover, it is

[1] Earman (1986).

[2] Albert (2000).

[3] This is so whether or not properties of such parts exhaust or fix the properties of the global state—I'm not relying on any supervenience theses.

arguable that this local determination is exactly what the original intuitive notion of a cause was supposed to capture.

The notion of a local event is a tricky one. At least, it must involve location within some spatiotemporally small area. Maybe it also has to be centered on some salient event, or on the location of the agent who characterises events as local. But I want something slightly stronger: something like 'epistemically local': macroscopically describable, medium sized, readily distinguishable from other events going on around it (hence discrete). One aspect of locality is that these events are easily subsumed within familiar and natural (to us) delineations of the categories of events. I take it to be a feature of our commonsense causal language that effects and causes are both local in this way.

Russell's other argument picks up on this conception of causation between local events. Consider some small local events c and e such that the occurrence of c determines the occurrence of e but not vice versa. Russell's argument is that these local events won't be the kind of things we typically take to be related by cause and effect. Consider a putative causal relationship between the firing of some gun F and the death of a victim D. These events are clearly too particular to get into the determination relation, because F can occur without D, if the bullet misses the victim, or if someone else intercepts the bullet, or if the bullet explodes harmlessly in mid air.[4] The problem with local determination is that there is always the chance of some *interference* from outside the local area at the time of the cause. To ensure that the cause guarantees the occurrence of the effect, we shall have to hold the cause to be a very large set of events, perhaps the whole past cone of potential causal influence on the effect. So if c really is a determinant of e, c will have to be incredibly more complex and larger than causes are typically taken to be.

To really determine e, we shall have to make sure that all possible interfering events don't occur, which will mean specifying the events which actually fill the location of those potential interferers. But this will involve some events that are intuitively not causes being counted as causes simply by virtue of their being in a potentially causally efficacious location. Indeed, we shall be unable to make a distinction as regards causal efficacy with respect to these events—in particular, we shall not be able to distinguish the genuine from the merely potential causes from a set of events each of which occupies a particular kind of location. If we can't distinguish a genuine cause from an actual non-cause that might have been a preventer, then we shall be unable to engage in goal directed activities which

[4] And *vice versa*, if the victim dies from another gunshot wound, as in preempted shootings.

depend on effectively bringing about certain states of affairs. This is partly because the set of events we shall have to manage or intervene on is far too large and inhomogeneous to deal with effectively. The real problem, however, is that we shall be misled into performing actions to bring about our goals that are not effective for achieving those goals. Cartwright (1979) emphasised this aspect of causation when she talked of effective versus ineffective strategies.

We can put the net result of these arguments as follows. Physics gives us a deterministic structure of the evolution of a system over time, so in physics the notion of a cause is trivial because it counts every past event as a cause. If we wish to apply the concept of causation to some spatiotemporally local situation, then we can only have determination if we are willing to abandon the role of causes in demarcating effective strategies for manipulating that situation, which is to say if we abandon the traditional concept of causation altogether. The genuine physical determination relation has nothing to do with the demand for effective strategies, and any addition of some other relation having to do with such strategies is redundant. If the notion is redundant in physics, it is dispensable from a properly scientific account of the world. Since fundamental science is the arbiter of genuine ontology, this relation of causation should be excised from our folk ontologies: it can't even be reduced to physics, let alone found within it.

2. Causation as partial dependence or default determination

It seems to me this is exactly the wrong conclusion to draw. Physics provides us with full determination which would trivialise the notion of causation. But it only trivialises that notion because we thought that causation required full determination. (Recall that it was the requirement of determination that made us include all those pseudo-causes.) Once we recognise that there are non-trivialising relations of partial dependence of one event on another that can serve as a respectable foundation for effective strategies, we should abandon the quest for fully determining causation. Events which are effective yet not foolproof at bringing about other events are legitimately called causes.

One very nice scheme for capturing a relation of non-determining dependence between local events is provided by the use of counterfactuals.[5] Lewis (1973a, 1979) provides a putative analysis of causal claims in terms of certain counterfactual conditionals: roughly, c directly causes *e* just in case the counter-

[5]It is especially nice because it preserves a kind of defeasible determination, in a way that (say) probabilistic analyses of causation do not.

factual $\neg c \ \Box\!\!\rightarrow \neg e$ ('if it hadn't been the case that c, then it wouldn't have been the case that *e*') is true, and is not a backtracking counterfactual. A backtracking counterfactual typically has *e* depending on c, but *e* preceding c, e.g. 'if it had not been the case that the glass was broken, then it wouldn't have been the case that I had smashed it earlier'.[6] I am supposing that excluding backtrackers captures one of the intuitive platitudes about causation and its relation to time—for me, the asymmetry of actual temporal dependence supports the temporal asymmetry of counterfactual dependence. Even if you think that counterfactual *analyses* of causation are flawed, there does seem to be some connexion between counterfactual reasoning and causation.

We might worry that counterfactual determination looks like full determination of *e* by c, especially in the marginal gloss. But that appearance is misleading. Ordinary conditionals obey the following inference rule: $\alpha \rightarrow \beta \vdash (\alpha \wedge \gamma) \rightarrow \beta$ (the rule of 'strengthening the antecedent'). But counterfactuals do not, and the way they don't is instructive in this context. Typically, were I not to have fired my gun, the victim wouldn't have died. But were I not to have fired my gun *and* someone else did, the victim would have died. But were I not to have fired my gun, and someone else did, but they missed, the victim wouldn't have died. And so on. Additional considerations and factors can alter the assessed counterfactual dependence. These additional considerations are typically potential events that we didn't consider in the initial attribution of causal effectiveness to the antecedent event.[7]

But there are some events which make no difference when added to the antecedent. Example: we don't think that were I not to have fired, and were some small event φ on Pluto to have occurred, then the victim would have died. This is true regardless of what φ is. Some events, no matter what their character, are not capable of affecting the counterfactual dependence between two other events. But had other facts about the event φ been different—had it for example been located on Earth—then some of the possible events φ could have altered the consequent—if it had been a firing of a laser ray for example. Again: if some alien had been located on earth and had decided to perform some action φ in the near vicin-

[6]One thought is that these backtracking counterfactuals involve evidential reasoning from symptoms to causes, though this cannot, without circularity, define backtracking.

[7]One situation where full determination does appear is if the antecedent and consequent events are global states of a system. Then $\Gamma_1 \ \Box\!\!\rightarrow \Gamma_2$ cannot be defeated by affixing additional events, for there aren't any. Russell's argument can be restated: given the fundamental physics we have, for any two global states such that $\Gamma_1 \ \Box\!\!\rightarrow \Gamma_2$, there is a true counterfactual $\Gamma_2 \ \Box\!\!\rightarrow \Gamma_1$. There is thus no asymmetry of determination, and hence no causal asymmetry.

ity of my firing and the victim's being shot, then some of his possible actions φ could have contributed to or detracted from my action's bringing about the victim's death, whether or not he actually performed one of that subclass of actions. Notice that if the alien fires his laser, he can overdetermine the victim's death—the victim is shot twice—yet from the point of view of whether the victim died or not, the result is the same.

I draw three morals from this little science fiction example. First, events are the wrong thing to be the causal relata, at least if we only consider actual events. Events are distinct if any aspect of them is different. So the victim's death is different if it is a laser-and-bullet death than if it is a mere bullet death. In fact, the identity of the event is quite fragile and quite sensitive to actual circumstances. Although some think that we should have a modally robust notion of event that isn't so sensitive, I'm quite happy to think that events are modally fragile in this way. But since I believe I would have played a part in causing a death regardless of the alien or not, I had better take the relata of the causal relation to be something else: maybe abstract event types, or better still *random variables*. A random variable is a function from possible events to numbers, where the numbers characterise certain features of the event. For example, I could characterise a certain class of events as the relevant class of events that might have resulted from my firing a gun, and use the random variable **Death** which takes value 1 if the victim died, and 0 otherwise (and is undefined on events which aren't relevant—where relevant will be spelled out below). This concept 'smooths out' variations in the event which do not give rise to a different value of the random variable; however, some other random variable might be sensitive to some of those variations.[8]

Second, what matters to the causal importance of a random variable is if at least one of its possible values can alter a counterfactual in which it features. For example, I take it that the alien's close spatiotemporal location to my shooting is a relevant class of events, because some of the events in that location involve him shooting a laser. That is, for some of the events A that fall into that class, a counterfactual $\neg F \wedge A \mathrel{\Box\!\!\rightarrow} D$ is true. But for this alien's activity to be *potentially* causally relevant, it doesn't matter that *actually* it was not.

Thirdly, suppose I am right about the causal significance of events which potentially have causal relevance to other events, due to the truth of certain counterfactuals. Then I need to have a way of ruling out all events which have a potential counterpart which has causal relevance. I think that this is largely a matter of con-

[8]This is what Field (2003) calls a 'fairly inexact variable', and is essentially the same notion of variable that appears in Hitchcock (2001), Pearl (2000), and Spirtes *et al.* (2000).

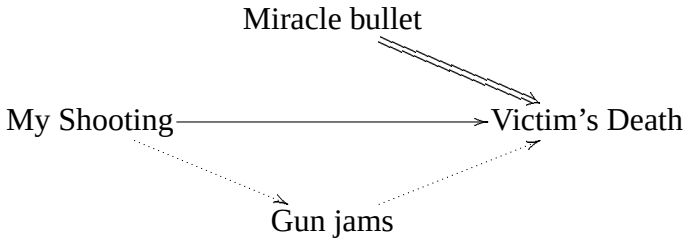

Figure 1: Types of variables: ordinary causal variables, such as my shooting; non-serious possibilities, typically not included, such as the miracle bullet; unlikely possibilities ruled out by default, such as the gun's jamming when fired.

textual salience, just as in traditional counterfactual accounts. But I also have a principled way of judging salience.

Let me expand. Standard counterfactual accounts have the feature that, if outlandish enough possibilities are considered, almost any robust modal counterfactual connection between events can be disrupted. If we think that A depends on B, consideration of a miraculous intervention, or time travel, can make B happen regardless of A. So too on this counterfactual account, and in two types of case:

(Type I) Some events are considered irrelevant, but if one were to consider them they would have a significant potential impact;

(Type II) Some events are relevant, but only very few of their 'values', or ways of occurring, have any significant impact.

Examples of the two types may be: (i) the victim's death depends on my shooting the gun at him, but not if a bullet would have spontaneously appeared in midair with the same trajectory as if I had shot it; (ii) the victim's death depends on my shooting the gun at the victim, but not if the bullet jams in the breach. The 'defaults' of Menzies (unpublished) can be used to account for causal judgements in cases of this second sort: if we have a probability distribution over the values of a random variable, then certain probable values can be used as defaults to make default causal judgements. (See Figure 1.)

What is going on with these counterfactual claims about possible events? I think that when we engage in counterfactual reasoning of this sort, what we are doing is constructing *ad hoc* scientific theories of the situation in question. We make a little model which has a few parameters, and the whole theory is the class of models compatible with the constraints we impose on the values of the parameters, including constraints imposed by the values of other parameters. Each

individual model is a possibility for the system in question. Which parameters are included will depend on our judgements about the physical nature of the situation in question, about the spatiotemporal connections (signals and the like) between the events which instantiate the variables, and on the purposes for which the model is being built. This last point is that models that support counterfactuals are often built to explain certain events, and that explanations are typically contrastive, set up against a context which demands that some salient variable or relation be included, and others may be neglected. One very important dimension of salience for my purposes will be the locality of the event with respect to us, for this will ensure that the variables we model are just those required to evade Russell's argument. Counterfactual reasoning depends additionally on which possibilities the model dictates we take seriously.[9] Which events are serious possibilities for us depends on model and context.[10] In line with Stalnaker (1984), I take context to be the information—including information about their capacities and goals—presumed available to the agents engaged in investigation or inquiry; the context set is the set of possible worlds compatible with this information. These are the 'possible situations' that the theory counsels us to consider when we wonder what might be.

I think the difference between the type I and type II cases above is simply that the type I cases are non-serious possibilities, and the type II cases are serious possibilities that are often not contextually salient because the disturbing values are highly unlikely. Hence my thought is that the type I claims depend on a contextual claim about model choice, i.e. which variables to include, and type II claims depend on the plausible default range of values that a variable can take (where plausible is to be cashed out in terms of frequency within the relevant class of models). For causal prediction, the default range of plausible values will constrain the application of the model. For instance, we will judge events which are unlikely values of their relevant variables to be not causally efficacious; though in thinking about actual past causation, we may well discard the plausible in favour of the actual values. But there is a choice of causal model which is made prior to any context of application of that model to a concrete situation. The contextual features of model selection, as opposed to model application, I am less sure what to say about. I think that facts about actual correlation and about actual spatiotemporal connectability between events that are phenomenologically salient will go a

[9]For example, in the Lewis-Stalnaker framework (Lewis, 1973b, Stalnaker, 1968), the 'seriousness' of a possibility depends on its similarity to actuality.

[10]The idea of a serious possibility comes from Levi (1980).

long way to explaining the choice of variable. One thing to note is that the truth of the appropriate counterfactuals will be a perfectly objective feature of a given set of variables, even if those variables are chosen for pragmatic reasons.

Let me deploy some new terminology to summarise the 3 morals about counterfactuals that I recently drew. Let the situation under consideration determine a contextually salient theory, specifying the set of random variables $\mathscr{V}$ that we will use to summarise the values of the events in question. The theory will encode certain counterfactual dependence claims. In particular, it will encode a pattern of mathematical dependence between parameter values for random variables. That is, it will give us facts of the form "the value of variable V_i depends on the values of variables $V_j \ldots V_k$." So, for example, the value of **Death** (yes (1) or no (0)) depends on the value of **Fires** (yes or no), and also on the value of **Alien** (fires, stands by, or throws himself in the way—so variables can be more than binary-valued). In general, however, not every variable will depend on every other: some will be *independent*.

The kinds of counterfactuals we take to be true will determine exactly how the dependency is cashed out. Call a variable V a *parent* of another variable U just in case there exists some assignment of fixed values to variables in the model such that the counterfactual "were V to have some different value v, then U would have some different value u" comes out true, and $V \neq U$.[11] Note that 'grandparent' variables are not parents: if V only acts on U through W, then the fact that W is fixed on some value will prevent the change in V from percolating through to U. This counterfactual (roughly equivalent to the notion of a *direct cause* in Woodward (2001)) allows us to construct *causal graphs* as follows. Take all the variables in $\mathscr{V}$, and put them at nodes of a graph. For each $V_i \in \mathscr{V}$, let $\mathscr{P}(V_i)$ represent its parents. For each variable V_i, draw an arrow from each $V_j \in \mathscr{P}(V_i)$ to the node V_i. We will end up with a graph something like Figure 2. This is a qualitative causal structure, and the parenthood relations are the most basic kind of counterfactual that should be considered in causal reasoning.[12]

This kind of structure will be quite familiar from the causal modelling framework introduced by Judea Pearl (2000) and the team of Spirtes, Glymour and Scheines (2000). The philosophical development I have given it here is basically

[11] Thanks to Charles Twardy and Chris Hitchcock for help with this formulation.

[12] If we have in addition a probability distribution over the values of the exogenous random variables (i.e.those with no arrow leading into them), and equations which express the numerical dependence of the values of a variable on the parameters and values of its parents, we can turn this qualitative causal structure into a quantitative causal model.

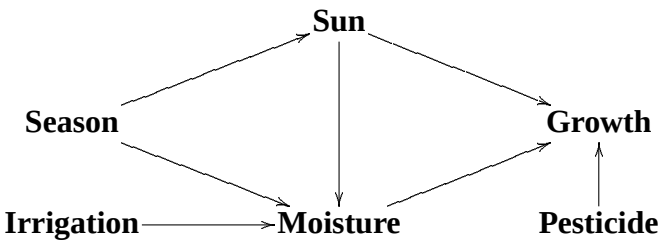

Figure 2: A sample causal model of plant growth and some of its factors.

that of Hitchcock (2001) and Woodward (2001), though I hope I have made it appear a very natural consequence of claims to which counterfactual analyses of causation are committed.[13]

Consider as a simple example the model of plant growth depicted in Figure 2. In this model, **Season** is binary: growing season or non-growing season, which in turn influences the amount of sunlight (**Sun**) and the rainfall (**Moisture**). But moisture is influenced by sunlight (causing evaporation) and whether the crop was irrigated or not (**Irrigation**).[14] Finally, whether **pesticide** is used or not also influences the final plant growth (**Growth**). Some of these variables are binary, and some are quantities (sunlight and rainfall). Different models may choose to represent irrigation by a volume of water, not by a binary variable. Similarly, growth may be modelled by a variable taking values of 'increased', 'decreased' or 'same'.

Interestingly, some variables can cause by 'omission' (i.e. not applying pesticide can negatively influence plant growth), and it may be noted that this framework gives an easy way for causes by omissions to work—at least, after the general problem of variable selection has been solved there is no special problem of causation by omission.[15] Of course, other irrelevant variables are also omissions of things that might have impacted, and it is the choice of variables to model along with judgements about salience and relevance that give content to our causal judgements.

[13]There may be an interesting connexion here with Lewis' recent account of causation as influence (2000). Lewis' account requires that for c to influence *e*, there must be a range of relevant alterations of c that are associated with relevant alterations of *e*. The concomitant variation of effect variables on cause variables in the above account may capture this, as well as yielding counterfactuals which give the influence a uniform treatment within the counterfactual framework.

[14]Irrigation and rainfall may themselves be correlated, so this model may not be perfect (since rainfall excludes irrigation typically).

[15]Thanks to Brett Sherman and Karen Bennett for help with this.

Several things are noteworthy about this approach to a counterfactual analysis of causation. Consider Figure 2 again. The values of the variables **Season** and **Irrigation** are not counterfactually dependent on anything (note that the restriction against backtracking counterfactuals is crucial here: if it were not, plausibly if the value of **Sun** were different, then the value of **Season** would have had to have been different too). But presumably these variables do depend on something—they are probably not the basic first events of the universe. To pretend we have isolated them, we can appeal to the contextual salience of local causes. Causal explanation has to stop somewhere, and if a certain condition on the parentless variables is satisfied, we should be prepared to stop with them. The simple condition is that parentless nodes should not be correlated amongst themselves. (I see this as a methodological condition on the construction of causal theories, not some a priori truth about systems of variables.) There is one difficult case: if X is counterfactually dependent on Y, and Y on X, neither through a backtracking counterfactual, then we should try to find another variable Z which is parent to both X and Y and screens off the counterfactual dependence. (However, there are some cases where such a variable does not exist, for example in standard explanations of the non-local correlations in Bell-type theorems in quantum mechanics (Butterfield, 1992). This is perhaps best modelled by simply keeping the two-way counterfactual dependence.)

Following naturally on from this, one can see how adding more variables can change the parental counterfactual dependencies by interpolating further intermediate causes, and by adding new parents. This can mean that contextual salience determines the causes of an event. So does the comprehensiveness of the underlying theory that supports the counterfactuals. This feature tends to support the idea that causation, as well as explanation, is often contrastive rather than absolute—it depends on the salient variables.[16]

Third thing to note: consider the counterfactual we used to evaluate the parenthood relation. It relies on a seemingly miraculous ability to vary the value of one variable while holding all others fixed on a certain value. This process has gone under the name of an *intervention* in the literature. In the graphical models, it can be modelled by severing the node from its parents and setting the value of the variable and other variables, effectively rendering some dependent variable independent. We are supposed to think of an intervention on X as encoded in a causal graph C that terminates in the fixing of some value for X independent of

[16]Hitchcock (1996).

other values of other variables not in C. [17, 18]

The notion of an intervention is strongly counterfactual, perhaps involving a set of models where the best system of laws might be different than the one governing the theory as a whole. One could think it was even a causal notion, and hence that the account must be circular. I think that since the variables representing the intervention itself are typically outside the model, the charge of circularity cannot be sustained. If we tried to make a global causal model, then serious difficulties would arise when trying to show how causal interventions can be naturalistically modelled in the same framework. One difficulty would arise when we are forced to make sense of the notion of an intervention in non-causal terms—once every event has a representative variable in the model, every dependence relation between events is already represented.[19] (I think this is exactly the point of Russell's arguments in this context!) In any case, this way of thinking about causal variables fits nicely with the manipulability account of causation that Woodward has recently defended.[20] But it can also be made to fit with other accounts (such as the mechanical conserved quantity view of Dowe (2000)) insofar as those accounts respect the modal claims that here I am suggesting are constitutive of causal dependence in general.

Causation will be a relation between general type variables that can be instantiated by many different particular events, and will apply in the single particular

[17]This isn't quite right, because expanding some graph D that contains X to also contain C would render C not perfectly efficacious in fixing X (C would no longer trump all other causal factors).

[18]The viability of the concept of an intervention seems to depend on the possibility of *modular* causal systems (Hausman and Woodward, 1999). These are systems where each variable in the system has some independent exogenous sufficient cause. Cartwright (2002) has argued that modularity generally fails. However, she seems to rely on the claim that actually non-modular systems are not possibly modular, and this claim seems false if one is willing to countenance counterfactual variation in the patterns of occurrence of instantiation of distinct variables.

[19]Similar remarks have been expressed by Pearl (2000, 350):

> If you wish to include the entire universe in the model, causality disappears because interventions disappear—the manipulator and the manipulated lose their distinction.

He also seems to think (Pearl, 2000, §4.1) that interventions are to be connected with human free will and the causal 'unconstrainedness' of human volition.

[20]In particular, the manipulability account of causation gives a nice way of motivating restrictions on the kinds of counterfactuals we consider to codify causal claims. The counterfactuals we have considered, about the results of holding variables fixed while varying others, are naturally thought of as modelled by hypothetical alterations of experimental situations. But some of these antecedent situations are *very* hypothetical, and require very distant worlds to evaluate them.

case only in a retrospective or derivative fashion. This is because we have identified causation with a feature of general models of phenomena rather than the actual happenings that a model can apply to. Once we see the actual values of the actual variables, then we can retrodict the actual effects various potential causes had, through the light of particular assumptions about what the natural or usual range of the exogenous variables is. We make a default causal model, to use the terminology of Menzies (unpublished), and by using a mixture of default assumptions and evidence about actual values, we can create a restricted version of the causal model that should give us the acceptable token causes for some particular case. This will be a model that only applies to the single case in question. The tokens of events in question are something like Lewis' *versions* of events: there may be a smoking event, which has many actual versions that instantiate it. A helpful way to view the events is as *coarse-graining* the space of possible versions of events.[21] They are the union of a lot of different ways of some description of an event being satisfied. The variables are mappings from events of some salient type to numbers, such that in some cases an event which is in the domain of the variable is not even a version of the characteristic event of the variable: the 0 value of the smoking event is at best a 'null' version of smoking.[22]

Finally, what is the meaning of the arrows? An arrow between X and Y does not mean simply 'X causes Y'. Rather, it means something like 'some values of X are causally relevant to the values of Y', where this causal relevance can be at times stimulatory, at times inhibitory, and so on. This seems to nicely feed into Hitchcock's (2003) recent claim that what is metaphysically primary is the multiplicity of causal connections, rather than some uniform notion of causation that is supposed to apply to all cases. Consider that some of the arrows might be purely inhibitory, some purely contributory, some a mix of both, and what constitutes the ground of the counterfactual claim might be very different in each case. Why think they can all be shoehorned into one neat causal metaphor like the 'cement of the universe'? One consequence might be that the distinction between

[21]I thank Chris Hitchcock, Helen Beebee, Graham McDonald and Huw Price for suggesting this way of viewing things to me.

[22]There is another sense of type here that may also have a role: this is that any particular event may be an instance of a lot of different coarse-grained events, so that an actual particular event e could be coarse event E_1 and also the distinct coarse event E_2. That is to say, particular events are subsumable under many different causal models, and judging in a particular case which is best is not altogether clear. There may not be a metaphysical fact of the matter, as will become clearer below.

a cause and an enabling condition will be harder to make in this framework.[23]

3. Determinism

Let us take stock. I think at this point we have a start on a framework for talking about counterfactual relations of dependence between convenient sets of local events. We can see that Russell's arguments depend in one sense on trying to give a Carnapian explication of the causal relation in terms of relations and events that physics makes available to us. We translate the talk of events into talk about the global states that constitute those events, and we translate talk of causal production into talk of laws of succession between microstates. In one sense all I want to suggest is that this explication of causation is a bad one. We aren't forced to map events onto global states; and we aren't forced to think that respectable science must give us invariable laws of succession. We have a perfectly viable framework for replicating causal talk as a certain pattern of counterfactual dependence between variables which describe the salient localised possible events.

But we may be left with a worry that the reconstruction I've given doesn't really address the spirit of Russell's arguments. I've argued that there is a certain pragmatically necessary emphasis in the folk notion of causation on local events and relations of counterfactual dependence between the theoretical representatives of such events. But in a sense I've evaded the original worry. I've shown that if we are content to give an explication of causation as it appears to us and in the special sciences, then Russell's claims don't straightforwardly attach. A defender of Russell's arguments will respond that given some plausible theses about the metaphysical primacy of physics and the claim that all facts supervene on facts about the global physical state, the problem with causation is that its emphasis on the local and the modal seems not to fit with the way we take the world genuinely to be.

More slowly, the argument runs like this.[24] If fundamental physics is to be the ultimate arbiter of ontology, then only the relations it posits are to be accepted in a correct scientific picture of the world. As Russell saw it, fundamental physics captures all that can truly be said of the world. Do relations of counterfactual dependence appear in fundamental physics? Only between the global states posited by the theory, as determined by the functional dependencies encoded in the state transition trajectories. And why are the global states important? Because they

[23]Thanks to Mark Schroeder on this point.

[24]Thanks to Gil Harman for help with this.

give us the answer to every question we could pose about the system, and this ambitious inclusiveness is exactly what we were looking for when we took physical facts to be those on which all other facts supervene. If we are to look only at parts of the global state, we are at best making up approximations that we need to propose in order to use our theories: we are no longer doing ontology.

Perhaps an analogy with objective chance is helpful. Almost everyone thinks that, if determinism is true, then all objective chances are 0 and 1. Why? Because if microstates 'really' constitute the actual instances of events of a probability space, and the theory which posits them has deterministic state transitions, then fixing an event will fix the future progress of the system through the state space and hence through the event space. It is at best an epistemic feature of macroscopic events that they seem to have non-trivial chances of coming about or failing to come about. Russell might say here that the notion of chance has no place in a scientifically respectable world view, but we don't need to go that far. All we need is that the place of chance in that world view is at best a redundant summary of other facts, and not a genuine part of fundamental ontology. Models which present the world as having objective chances are false.

The exactly similar response I imagine being made by a defender of Russell with regard to causation. What we've shown is that if we take into account some quite natural epistemic limitations on human agents, we are forced to admit that we will be ignorant of the microstates that really constitute the world around us. We thus have to resort to less epistemically demanding notions in order to get on in the world. But these epistemically less demanding notions are simple but in fact false descriptions of the underlying world. The facts of fundamental physics simply are not structured in the way they would have to be for causal models to even approximate the underlying reality.

Indeed, the very features of causation that we've discussed bear this out. Causation is context-dependent: it is sensitive to which events or variables are included in the model, and some think it is relative to default values for the variables also. Causation is partial and local. If try to expand causation into a global and context-insensitive (yet still categorical) notion, by adding in *all* the variables, then it seems to evaporate. If it holds at all between global states, then it holds trivially between every past event and every future event. And if it is not expanded, then it is at best a poor approximation to what physics tells us more satisfactorily.

At least probability can have some kind of robust existence. If Reichenbachian frequentists are right, then one has chances everywhere one has a sequence of outcomes that can be partitioned and whose frequencies converge. Chance then is merely a summary of what physics can tell us in more detail. But causation

misrepresents genuine physical relations.

The really curious part is how all this connects up with determinism, the notion that causation was originally supposed to go hand in glove with. We have now the curious situation that non-trivial causation and determinism are at odds. Suppes (1970) argued that causation re-appeared in fundamental science just when probabilistic indeterminism at the microlevel became apparent. If there is no further fact that fixes the value of some variable X, then different values that X could take in some future evolution of the system will make a difference to further later characteristics of the evolution of the system. But if determinism is true, then there is always some further fact that fixes the actual value of any variable on every actual event once some set of occurrent events is fixed. X can't take different values, hence can't make a difference. Of course, if it had been different, then everything would have been different. There is no sense in holding some things fixed and letting other things vary if the determination relation holds between whole states.

4. Perspectivalism

We could evade this neo-Russellian argument if we could show that the positing of these epistemically convenient and local relations (like the causal relation between variables) is not undermined by the fact that fundamental physics doesn't include them. There are a number of ways we could do this, some less plausible than others. One quite plausible way goes in two steps. First, see that the reason theories get accepted isn't always that those theories represent 'fundamental reality'—sometimes, for example, what governs acceptance is the acceptor's perspective on the issues under consideration. Second, see that from within a context where one accepts a certain theory, ontological questions about the status of the entities and relations in the models of the theory are settled by that theory itself—the content of that theory is the standard for correctness.[25]

Picking up on this idea of a perspective, Huw Price has defended a *perspectival* approach to modal notions in science that I think can help us escape the neo-Russellian argument for the elimination of causation (that because there is no reduction of causation to the concepts of fundamental physics, causation should be eliminated).[26] The position I will defend applies a similar kind of perspectival-

[25]Carnap's distinction between 'internal' and 'external' questions is similar to this second point (Carnap, 1956), but here I talk in terms of entities and models rather than rules for the manipulation of expressions involving novel or doubtful terms.

[26]Particularly Price (unpublished); see also Price (1992, 1996, 2001).

ism to causation, and will show why the approach to causation outlined above is particularly germane to the pragmatist.[27]

The resurgent form of Russell's argument that we sketched above rests on a broadly realist and imperialist approach to the interpretation of scientific theories. Scientific theories purport to be true and accurate *representations* of the way the world is; and the genuine facts are entirely captured by the pronouncements of some fundamental theory—in this case, fundamental physics.[28]

To resist this argument we need to resist this approach. The easiest motivation for resisting it is to see just what it misses out. Firstly, it fails to account for the fact that scientific theories are not merely or even primarily favoured for their representation-derived features (like truth), and the respects for which they are favoured (like explanatoriness) can undermine these features. Secondly, it fails to account for the fact that the agents who use scientific theories have a particular position in the world that influences the kinds of theories they propose and adopt.

This second point, suitably reflected on, gives us a start on a better understanding of science. We begin by recognising that we are *agents*, trying to make our way in the world. As such, we have a perspective on reality that reflects our status as agents, hence as having goals and projects, capacities to accomplish some of these and the resources to deliberate about how to best achieve them. It is no surprise therefore that the concepts that we bring to bear in dealing with the world reflect this goal-directedness and our consequent interest in effective strategies. This is particularly apparent in the case of causation, which gives us a nice way of systematising and approaching how we bring about effects and intervene in systems to get them to behave in ways that are fruitful for us.

It is also no surprise that the kinds of things we take effective strategies to bring about are salient and local events, and the interventions which effectively bring them about are also salient local events. We are limited in the scope of our activity and hampered by our merely partial control of its circumstances. Understanding the causal relation is a matter of understanding how agents like us should wish to systematise their views on how to characterise effective goal-directed activity in a world where their agency is constrained in roughly the same ways that ours is.

This family of capacities that we possess gives us a perspective from which

[27] One difference between my view and Price's is his emphasis on expressivism about discourse, whereas I prefer to think in terms of pragmatism about theoretical content. But the views are similar enough to deserve the same name.

[28] I will ignore the possibility of denying the claim that there is any fundamental level at all, recently defended by Schaffer (2003).

we engage with the world. From this perspective, the world presents relations of causal relevance between events—that is, the models we correctly adopt of the phenomena as they appear in our perspective posit causal relations between the variables. To creatures who are not concerned with their effective agency, or who are of such extent and duration that they need not be worried about locality, or who are of such cognitive capacity that they need not worry about salience, the world will look dramatically different: by which I mean that the models they choose to adopt on the basis of their experience will be radically different than the models we adopt.

We can posit some other perspective, a set of concepts from fundamental physics for example, that will abstract away from the logical dependence of our concepts on our limited purview. But we can't make the mistake of thinking that this other family of concepts will replace ours, or that they will make ours illegitimate. Our concepts are legitimated by our practices and are valid within the framework of those practices. That is, the standard for correct application of those concepts derives from our practices and capacities, and when we engage in those practices we are completely entitled to use those concepts—regardless of what other practices might entail about them. Determinism and time symmetry, concepts from fundamental physics, derive their legitimacy in turn from a particular perspective, which characteristically attends only to global invariant features of its subject matter. This perspective might not validate the concept of causation, but no further perspective can adjudicate on the relative merits of these families of concepts, except as they are appropriate or inappropriate for various tasks.

The realist may respond that relying solely on the pragmatic necessities of our own agency to support the concept of causation is perfectly compatible with there being no genuine metaphysical relation of causation. (The realist is apt to think that, in failing to attend to global and invariant features, a causal perspective is defective and hence not able to give a trustworthy account of genuine ontology.) To respond to this worry we need to consider how to understand the talk of adopting a theory that we used two paragraphs ago: that is, adopting a particular set of models of a phenomenon as a guide to that phenomenon. Our strategy will be to show that a correct understanding of acceptance will give easily met conditions under which the concept of causation is a perfectly legitimate part of an accepted theory, and hence not suitable for wholesale elimination.

An appealing way, once we have introduced the users of a scientific theory into the picture, is to understand the adoption of a scientific theory naturalistically. That is, we have an account of the agents in question, and their requirements of the theories that they use to guide themselves through the world. When combined

with an account of the way the world genuinely is, this need not mean that the only goal when adopting a theory is that it be true. Rather, theories are adopted because they conduce the ends the agents have. A toy example might be this: agents who are extremely cognitively limited might adopt the theory that best allows them to get along in the world, which means a very simple theory they can actually use, and not necessarily one that captures the world in all its complexity (in some sense we are such agents).

In our case, we need a class of models that best fits with our particular cognitive and epistemic features. The natural way to understand this is to look at models of us and our use of theories: what Price calls a 'meta-model'. I propose one particular meta-model. This model contains creatures who suppose, for the purposes of effectively getting along in the world, certain theories. They make these suppositions in the context of diligent inquiry into the world, but that context has two features: first, it presupposes that the appropriate model lies somewhere within a set of possibilities which include the relations deployed by those creatures; and second, consideration of particular projects and tasks can shift the context by altering the legitimate presuppositions of the inquiry.[29] But certain features of the context are, though variable amongst different kinds of agents, fixed for each particular kind. These will be the psychological features of the agents, their particular characteristic ways of approaching the world.

The meta-model describes what it is for such creatures to suppose a particular theory in a context—it is for that context to leave open certain possibilities, the very possibilities that the theory has as models, and for those possible events to be taken as the relevant 'outcomes' which in turn are inputs to a process of rational deliberation. In this case, the agents act as if the theory they accept is true; from their perspective, the theory provides an all-embracing background for their deliberation.

The crucial point for our purposes is that psychological and contingent features of a kind of agent enters into the specification of the context. Therefore, theories which leave out features that are present in all the possibilities consistent with that context are poor candidates for acceptance.

Let us make what I take to be a very weak supposition: that scientific theories are used to give explanations, make unsurprising already observed phenomena and guide expectation about unobserved phenomena, in addition to describing the world. Additionally, simple and comprehensive theories are often more acceptable. Each of these virtues are not necessarily correlated with the likelihood of

[29] Stalnaker (1984).

truth of a theory. If they nevertheless play a role in governing the acceptability of a theory, which it is apparent that they do, then acceptability is not predicated solely or even primarily upon accurate representation of the world. Rather, acceptance of a theory will depend on its satisfaction of these pragmatic and theoretical virtues, and how well it fits with the background context settled by our particular psychological stance as agents. This is particularly apparent if accurate representation is regarded as independent of the psychological features and pragmatic goals of the users of those representations: for then our particular context will almost inevitably yield additional structure that our theories possess that do not appear in the 'agent independent' representation.

An accepted theory governs the beliefs and actions of the agents who accept that theory in exactly the same way as if they regarded that theory as truly representing: one cannot accept a theory as non-representational. But a moment's reflection on the meta-model of the phenomena of these agents' acceptance is sufficient to show that any accepted theory is treated as true only in order to facilitate the agents' use of the theories in question. That is not to say, however, that there aren't contexts where representation is primary, for there are. The context of fundamental physics may be one; and so may be the context in which we inquire into the naturalistic position of human agents within the physical world, as physical entities with certain psychological features. So this position is not anti-realist, if that means to deny the possibility of scientific inquiry having the primary aim of truth. This position does claim, contrary to scientific realism, that being an element of the 'one true theory' is not necessary for being an acceptable theoretical element in contexts of inquiry which can justify our family of concepts, namely those contexts we are inescapably embedded in which involve those concepts.

One particularly interesting feature of this contextual meta-model of acceptance is that for certain creatures, a purely representational theory would be both impossible and inadvisable to accept. Impossible, because the serious open possibilities for such creatures (as partially determined by their psychological endowment) do not include the possibility described by the representational model. Inadvisable, because the kind of conceptual reform that it would require in order to get such creatures into a state where acceptance of that theory was possible would be of such magnitude and scope that it would render them unable to successfully make their way in the world. A plausible meta-model of human inquirers has it that we are such creatures, at least with respect to the notions of chance and causation. Without either notion, we would be rendered incapable of deliberating and acting; but of course we could never seriously entertain the possibility of jettisoning them either.

In short: our cognitive abilities and pragmatic goals set a context for the theories we accept which demands those theories be sensitive to, and do not undermine, the perspective we have on the world: the way the world seems to creatures like us.[30] This context makes legitimate our adoption of the ontologies those theories encode, since we can neither escape it nor function effectively without it. We act as if our adopted theories were representational, since that is a trivial consequence of what it means to accept (and is a rational strategy for us to adopt with respect those theories we use to guide our activity). However, in principle we could ascend to a meta-model and discern that the model actually is one that best satisfies our aims and goals, and may not be the best representation. Hence correct representation cannot be the sole arbiter of ontology—there remains the possibility that the item in question appears legitimately and irreducibly in another accepted theoretical framework.

An example of this might help. Often, given our epistemic limitations, there is a need to *coarse-grain* the set of states of a system: to summarise a complex state by means of a probability distribution over a set of coarser properties.[31] I think that, given a context where a theory involving a particular coarse-graining is accepted, it is correct to claim that if the theory works and provides explanations and predictions, then what it proposes should be treated as existing. In particular, sometimes when we consider quasi-isolated systems with a particular coarse-graining of the events of interest, the causal modelling framework described above can have this status. Since this framework yields a number of causal relations, they exist. Whether or not determinism reigns, for the kind of system in question this approach is the best we have. Now, it just so happens that around here, quasi-isolated systems are instantiated by deterministic systems. But that doesn't mean that there is no possibility that they might not be; nor does that impugn the fact that approaching these deterministic systems in this way is methodologically sounder than the alternative reductionist approach. That is to say, given the current context, there is a serious possibility that the acceptable theories are indeterministic, given what they claim about macroscopic quasi-isolated systems. To pretend that there is no such possibility is to make a serious methodological error given the particular perspective one operates within.[32]

[30]The way the world seems to us will include the entities and relations we project onto the world in virtue of having cognitive capacities that are sensitive to particular patterns of correlation of salient events.

[31]For a similar characterisation, see Strevens (2003).

[32]This example was suggested to me by Chris Hitchcock.

In a sense, this is a fictionalist approach to scientific theorising. Instead an 'according to the fiction f' operator, we have an 'according to the theory t' operator that encloses all the direct assertions of agents in particular contexts. From within a contextual perspective, just as from within a fiction, the world seems just as the theory or fiction describes. When modelling creatures as users of a particular theory, what we do is answer the question why these creatures should find it congenial to adopt the fiction in question. However, in denying the necessarily representational character of scientific theories (when taken in isolation from a context of acceptance), we are at odds with the fictionalist, since we can't avail ourselves of any context-free non-fictionalist approach to any theory. A model, in the absence of a context for acceptance, cannot be regarded as describing anything, for we have not fixed anything for it to attempt to match up to.

Price puts it nicely when he characterises the whole approach as republican (in contrast to Russell's 'monarchy or anarchy' ultimatum): once we see that acceptance of a model can be motivated partly by features of the agents who accept, then the dichotomy between objectivism and eliminativism can be seen as incomplete. That a theory deploying a relation or object is acceptable in some contexts and not in others indicates that the relation or object is not 'real' in some context-independent sense. However, that does not indicate that the relation or object must be eliminated, since by hypothesis it was acceptable in some context. Nor can it be reduced without remainder to some context-independent notion, for the same reason. The entity remains irreducibly perspectival, and yet remains a viable component of a complete ontology. Agents needn't be given authority to accept a theory and its constituents by fundamental reality: lacking such an authority they can provide it for themselves if their character and the available evidence is suitable. Indeed, in some cases inescapable features of their situation demand a commitment to entities that do not feature in fundamental reality.

5. Causal Counterfactuals and Deliberation

This has been quite general so far, but now I want to apply this perspectival framework to the particular view of causation I defended above: that causal claims are expressions of certain patterns of counterfactual dependence.

I've claimed that the feature of our perspective that causal claims latch onto is our status as deliberating agents. In this, I follow Ramsey:

> . . . from the situation when we are deliberating seems to arise the general difference of cause and effect. We are then engaged not on dis-

> interested knowledge or classification... but on tracing the different consequences of our possible actions... (Ramsey, 1929 [1990], 158)

We are trying to decide what to do in a world where our epistemic access is inherently limited to a local area. Even if there is no objective relation of causation, there would be a use for one if, by using it, agents could mark effective ways of achieving their goals. This would mean that causal claims had to provide an effective encoding for information about independence of one event from another (to tell us what wouldn't be effective). Causal claims would also need to give us a quick way of seeing just which events depended on which others; and how the patterns of dependence go. It should be clear that the causal networks of variables linked by counterfactual dependencies satisfy both of these desiderata.

Other notions of causation, though they too might be able to be ensconced in a broadly perspectival framework, are not going to have as close a connexion to deliberation as the counterfactual framework. In some sense, hypothetical reasoning as represented by counterfactuals is constitutive of rational deliberation. When considering what to do, we should maximise our expected utility. That is, we should consider the possible outcomes that might ensue given our act and then weight them by their subjective probability— that is, in our framework, what kind of distribution over the variables a particular interventional setting of some particular variable would induce—and by their subjective value. Other accounts of causation, like conserved quantity theories, have a much harder time getting this modal hypothetical aspect of deliberation out of their actualist stance towards the propagation of causal influence. Maybe they deal better on some other criteria, but as it stands the deliberative aspect of our agency seems like the most plausible feature of our perspective to ground the utility of causal models, and this account of causation neatly latches onto this ground.

6. Conclusion

I think that we should resist the Russellian conclusion. Insofar as we need to put constraints on our notion of causation, we have seen that it does not have a natural home in a deterministic global microphysics.

However, the concept of causation can be naturally reconstructed in contexts where the events we deal with are local and epistemically convenient. Indeed, in such contexts it has decided advantages over the reductionist aspirations of fundamental physics. This is especially apparent once the demand that a complete

ontology must only reflect perspective-independent reality is given up as inconsistent with our status as inquiring agents.

Draft of 23 October, 2003.

Bibliography

Albert, David Z. (2000), *Time and Chance*. Cambridge, MA: Harvard University Press.

Butterfield, Jeremy (1992), "Bell's Theorem: What it Takes". *British Journal for the Philosophy of Science*, vol. 43: pp. 41–83.

Carnap, Rudolf (1956), "Empiricism, Semantics and Ontology". In *Meaning and Necessity*, pp. 205–21, Chicago: University of Chicago Press.

Cartwright, Nancy (1979), "Causal Laws and Effective Strategies". *Noûs*, vol. 13: pp. 419–37.

——— (2002), "Against Modularity, the Causal Markov Condition, and Any Link Between the Two: Comments on Hausman and Woodward". *British Journal for the Philosophy of Science*, vol. 53: pp. 411–53.

Dowe, Phil (2000), *Physical Causation*. Cambridge: Cambridge University Press.

Earman, John (1986), *A Primer on Determinism*. Dordrecht: D. Reidel.

Field, Hartry (2003), "Causation in a Physical World". In Michael J. Loux and Dean Zimmerman (eds.), *Oxford Handbook of Metaphysics*, Oxford: Oxford University Press.

Galavotti, Maria Carla, Suppes, Patrick and Costantini, Domenico (eds.) (2001), *Stochastic Causality*. Stanford: CSLI Publications.

Hausman, Daniel M. and Woodward, James (1999), "Independence, Invariance and the Causal Markov Condition". *British Journal for the Philosophy of Science*, vol. 50: pp. 521–83.

Hitchcock, Christopher (1996), "The Role of Contrast in Causal and Explanatory Claims". *Synthese*, vol. 107: pp. 395–419.

——— (2001), “The Intransitivity of Causation Revealed in Equations and Graphs”. *Journal of Philosophy*, vol. 98: pp. 273–99.

——— (2003), “Of Humean Bondage”. *British Journal for the Philosophy of Science*, vol. 54: pp. 1–25.

Levi, Isaac (1980), *The Enterprise of Knowledge*. Cambridge, MA: MIT Press.

Lewis, David (1973a), “Causation”. In Lewis (1986), pp. 159–213.

——— (1973b), *Counterfactuals*. Oxford: Blackwell.

——— (1979), “Counterfactual Dependence and Time’s Arrow”. In Lewis (1986), pp. 32–66.

——— (1986), *Philosophical Papers*, vol. 2. Oxford: Oxford University Press.

——— (2000), “Causation as Influence”. *Journal of Philosophy*, vol. 97: pp. 182–97.

Menzies, Peter (unpublished), “Causation in Context”. Presentation at *Causal Republicanism*, Centre for Time, University of Sydney, 14/07/03.

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

Price, Huw (1992), “Agency and Causal Asymmetry”. *Mind*, vol. 101: pp. 501–20.

——— (1996), *Time’s Arrow and Archimedes’ Point*. Oxford: Oxford University Press.

——— (2001), “Causation in the Special Sciences: the Case for Pragmatism”. In Galavotti *et al.* (2001), pp. 103–21.

——— (unpublished), “Russell’s Lost Republic”. Presentation at *Causal Republicanism*, Centre for Time, University of Sydney, 14/07/03.

Ramsey, F. P. (1929 [1990]), “General Propositions and Causality”. In *Philosophical Papers*, pp. 145–63, Cambridge, UK: Cambridge University Press.

Russell, Bertrand (1913 [1963]), “On the Notion of Cause”. In *Mysticism and Logic*, pp. 132–51, London: George Allen and Unwin.

Schaffer, Jonathan (2003), "Is There a Fundamental Level?" *Noûs*, vol. 37: pp. 498–517.

Spirtes, Peter, Glymour, Clark and Scheines, Richard (2000), *Causation, Prediction and Search*. Cambridge, MA: MIT Press, 2 ed.

Stalnaker, Robert C. (1968), "A Theory of Conditionals". In Nicholas Rescher (ed.), *Studies in Logical Theory*, pp. 98–112, Oxford: Blackwell.

——— (1984), *Inquiry*. Cambridge, MA: MIT Press.

Strevens, Michael (2003), *Bigger than Chaos: Understanding Complexity through Probability*. Cambridge: Harvard University Press.

Suppes, Patrick (1970), *A Probabilistic Theory of Causality*. Amsterdam: North-Holland.

Woodward, James (2001), "Probabilistic Causality, Direct Causes and Counterfactual Dependence". In Galavotti *et al.* (2001), pp. 39–63.