Common Causes and The Direction of Causation[*]

Brad Weslake[†]

28 July 2005

[†]Centre for Time
Department of Philosophy
Main Quad A14
University of Sydney
NSW 2006
Australia
brad.weslake@arts.usyd.edu.au
http://www.usyd.edu.au/time/weslake/

Abstract

Is the common cause principle merely one of a set of useful heuristics for discovering causal relations, or is it rather a piece of heavy duty metaphysics, capable of grounding the direction of causation itself? Since the principle was introduced in Reichenbach's groundbreaking work *The Direction of Time* (1956), there have been a series of attempts to pursue the latter program—to take the probabilistic relationships constitutive of the principle of the common cause and use them to ground the direction of causation. These attempts have not all explicitly appealed to the principle as originally formulated; it has also appeared in the guise of independence conditions, counterfactual overdetermination, and, in the causal modelling literature, as the causal markov condition. In this paper, I identify a set of difficulties for grounding the asymmetry of causation on the principle and its descendents. The first difficulty, concerning what I call the *vertical placement* of causation, consists of a tension between considerations that drive towards the macroscopic scale, and considerations that drive towards the microscopic scale—the worry is that these considerations cannot both be comfortably accommodated. The second difficulty consists of a novel potential counterexample to the principle based on the familiar Einstein Podolsky Rosen (EPR) cases in quantum mechanics.

Contents

I. INTRODUCTION: WHY DO METAPHYSICS?

In *The Direction of Time* (1956), Hans Reichenbach developed a theory of causation that included two elements—now unremarkable—that at the time marked significant departures from philosophical orthodoxy. Firstly, he held causation to be probabilistic[1]. Secondly, he attempted to give an explanation of causal asymmetry in terms independent of temporal order. Central in his account of causal asymmetry was what he termed *the principle of the common cause*—informally, if "an improbable coincidence has occurred, there must exist a common cause" (p. 157). The temporal orientation of causation is explained by the contingent fact that common causes tend to lie in (what we call) the past rather than (what we call) the future of the coincident events in question.

For Reichenbach the principle of the common cause was neither autonomous nor fundamental, being derived from the second law of thermodynamics and his hypothesis of the branch structure of thermodynamic systems. However in the subsequent literature the principle—or at least the key insight it embodies—has become increasingly detached from this context and developed into independent reductive grounds for the direction of causation. In addition to direct heirs of Reichenbach such as Dowe (2000), the common cause principle has appeared in the guise of independence conditions (Ehring, 1982; Papineau, 1985a,b, 1993; Hausman, 1998), counterfactual overdetermination (Lewis, 1979, 1986c), and, in the causal modelling literature, as the causal markov condition (Pearl, 2000; Spirtes *et al.*, 2000). Of course, those developing the causal modelling framework do not take themselves to be providing an account of the asymmetry of causation. For a start, they do not standardly take themselves to be pursuing reductive projects in the first place. Moreover, if you ask a causal modeller about the place of the direction of causation in their work, they will tell you that temporal orientation is simply *assumed*[2]. But—science for the scientists, philosophy for the philosophers—there have been a number of philosophers who take it that the causal modelling framework provides reductive grounds for the direction of causation after all (see Papineau, 1993; Field, 2003).

Why should we not dismiss these latter views on the grounds that they are unfaithful to the very theories on which they wish to hang their metaphysics? Since

---

[1]While Good (1961a,b) is a much ignored early proponent of probabilistic causation, my own ignorance is for present purposes justified, as he assumes the temporal orientation of causation from the outset. Suppes (1970), a much more widely known treatment, does likewise.

[2]There are at least two reasons why this is the case. Firstly, the various causal discovery algorithms on offer typically deliver a set of compatible causal models (a so-called *markov equivalence class*) rather than a unique causal model for any set of probabilistic data—and so additional information is required to select the correct model. Secondly, for any reasonably complex system the algorithmic search space will be extremely large—and so again, any available information which could reduce the search space will normally be employed. Temporal order is an obvious candidate in both cases. On the face of it, the latter poses less of a problem for the metaphysical reductionist, since the role of temporal information can here be reasonably construed as pragmatic or heuristic. Such a strategy sits less easily with the former, however; this will be discussed further in what follows.

our best frameworks for causal modelling simply assume a direction of causation, shouldn't we expect that thinking about the temporal asymmetry of causation in light of these frameworks will be at best unfruitful, at worst circular? I think this is too hasty. Indeed, I think that attention to the issues here is of both independent and practical interest, for at least two reasons.

Firstly, getting clear about the situations under which one would be prepared to accept the existence of backwards causation provides a means by which the concept (or concepts) of causation, so interwoven with other temporal and modal concepts, can be isolated and clarified. The questions here are those such as, which temporal facts can be varied without varying causal facts? Likewise, which modal facts can be varied without varying causal facts? Much recent debate in the philosophical literature on causation revolves around precisely these sorts of questions. Secondly, and most importantly for the purposes of this paper, is the converse of the preceding point. This is that taking the various methods of discovering causal relations on offer, and imagining the features which they detect to exist in the reverse temporal direction, allows the exploration of whether these features can be said to be *constitutive* of the causal relation or just *defeasible heuristics* for the causal relation. In the remainder of the paper I examine this latter question with respect to the principle of the common cause.

Before I proceed, however, some deflationary remarks are in order for those who suspect that the method I have just sketched sounds a little too much like conceptual analysis. It has become increasingly frequent in the literature on causation to claim that what is being attempted is *not* conceptual analysis but rather the *a posteriori* identification of what causation is, in the physical world. Such an approach is explicitly endorsed by Menzies (1996), Kistler (1999), Dowe (2000), Steel (2004), and, in order to defend his theory against a series of objections formulated in the way I have just described, by Papineau (1988). The idea here is to deny the link between conceivability and possibility; here is Papineau (1988, p. 524):

> I don't claim that such cases [conceivable counterexamples] are conceptually impossible. My claim is simply that my theory identifies the real nature of the cause-effect relationship.

As a general point, this is undoubtedly true—I can imagine myself flying, and yet I cannot fly. The general point here is also, I take it, applicable to various of the more metaphysically speculative thought experiments in the literature (the *locus classicus* here is Tooley, 1990)[3]. Nevertheless, the reply is only available for those counterexamples which rely on grounds weaker than the theory being criticised. In general, we should not expect this to be clear-cut; and in particular cases, we can be confident that the burden of proof is on the proposed theory. That is, we can, with the *a posteriori* reductionist, deny the link between conceivability and possibility,

---

[3]For example, why should we concede that the conceivability of simple worlds with uninstantiated causal laws as brute simples entails that in our world, causal laws can not be reduced?

and yet still do thought experiments to assess proposed reductions—*if we have good independent reasons for thinking the situations we have described nomologically possible.* Nobody objected to the Einstein *et al.* (1935) thought experiment concerning the completeness of quantum mechanics; and rightly so, since there were excellent theoretical grounds for believing it to be possible. Likewise, in this paper I will not be offering anything controversial by way of thought experiment—indeed, I will just be describing what goes on in certain cases of actual scientific and everyday practice. And so this loophole will not be available for those theories I take as my targets.

II. THE COMMON CAUSE PRINCIPLE

At first glance, the informal expression of the principle of the common cause given by Reichenbach above ("if an improbable coincidence has occurred, there must exist a common cause") cannot be what is intended. Reichenbach is surely not committing himself to the impossibility of one-off, purely chancy coincidences. Rather, it is repeated incidences of some *prima facie* coincidence that call for explanation—the improbable correlation of two (or more) events over time. This is made clear in the formal probabilistic definition of the principle given by Reichenbach (1956, pp. 157-167). He writes (p. 163):

> If coincidences of two events *A* and *B* occur more frequently than would correspond to their independent occurrence, that is, if the events satisfy relation (1), then there exists a common cause *C* for these events such that the fork *ACB* is conjunctive, that is, satisfies relations (5)-(8).

The relations being:

(1) $P(A\&B) > P(A) \times P(B)$

(5) $P(A\&B|C) = P(A|C) \times P(B|C)$

(6) $P(A\&B|\neg C) = P(A|\neg C) \times P(B|\neg C)$

(7) $P(A|C) > P(A|\neg C)$

(8) $P(B|C) > P(B|\neg C)$

That is, conditionalising on the common cause, and on the absence of the common cause, renders the effects independent (this is often referred to as the common cause *screening off* the effects); and the common cause raises the probability of both effects individually. In the language of independence, effects of a common cause are correlated, but causes of a common effect are not. In the language of overdetermination, there are many distinct effects which individually determine the occurrence of some cause, but there is only one cause that individually determines any effect. Obviously, these are rough and ready formulations of the related principles, and there are significant differences in the way they have been developed into complete theories of causal asymmetry. It is not the place of this paper to

5

explore these differences in any detail[4]. Rather, in what follows I will restrict discussion to the principle as given above, and indicate where the discussion carries over to the neighbouring theories.

The principle as stated is not yet sufficient for an analysis of causal asymmetry. To see this, consider any device which records correlated events. To modify an example from Reichenbach (p. 158), suppose two nearby geysers (*A* and *B*) spout irregularly, but (nearly always[5]) in unison—and suppose some geyser-enthusiast sits nearby, recording every dual spout with a tick in her notebook (*D*) (see Figure 1). The probabilities of each geyser spouting and the ticks being recorded in the notebook satisfy (1) and (5)–(8) (p. 162), and yet the tick, which is the screening off event, occurs later than the correlated geysers, which are the events screened off (*A* and *B*). The events are said, in the jargon, to constitute a conjunctive fork *open to the past*. Reichenbach is thus led to expand on the principle, by claiming that conjunctive forks open to the past are always matched by conjunctive forks open to the future, while conjunctive forks open to the future are not likewise matched. Here, for example, the event establishing the fork open to the future would be the increased water pressure in the reservoir responsible for the geyser spouts (*C*). Reichenbach refers to cases where correlated events are screened off on both temporal ends as *double forks* (p. 159); they are also sometimes called *closed forks*. The common cause principle thus becomes: all open forks are open to the future.

Note that in order for the arrows to be drawn onto closed forks, there must be a predominance of forks open to one direction rather than another; otherwise the temporally reversed interpretation would be equally justified. So, in Figure 1, it is the (hypothetical) predominance of forks open to the right side of the diagram that allows us both to say that the direction of causation runs left to right, and (therefore) that the arrows have been drawn correctly onto the closed fork[6]. Note also that if the common cause principle is constitutive of the direction of causation, we should be licensed by the existence of forks open in the opposite direction to infer the existence of backwards causation.

As Reichenbach notes, this is an "indirect" solution (p. 162), since it makes causal asymmetry a function of factors extrinsic to the events in question—and this is one place where the cluster of theories referred to above diverge. Hausman, Lewis and Ehring seek to keep the asymmetry intrinsic, while Dowe, Papineau and Field—after Reichenbach—opt for an extrinsic solution. Each have their problems, and I will return to the issues this introduces in Section III. For the moment, note that the correlation given by (1) is crucial here. Any device which records coin-

---

[4]Unfortunately, I know of no comprehensive survey of the differences and similarities between the theories, though Hausman (1998) provides some detailed criticism of each of the others.

[5]When I presented this paper at Konstanz without this qualification, Chris Hitchcock and Iain Martel were quick to point out that Reichenbach explicitly disallowed probabilities of unity, since this prevents the disambiguation of causal asymmetry by the principle of the common cause. The remainder of the paper should be read with this qualification implicit—it does not alter the structure of the argument.

[6]In fact, Reichenbach appealed to networks of probabilistically related events rather than global predominance. I will return to this point in what follows.
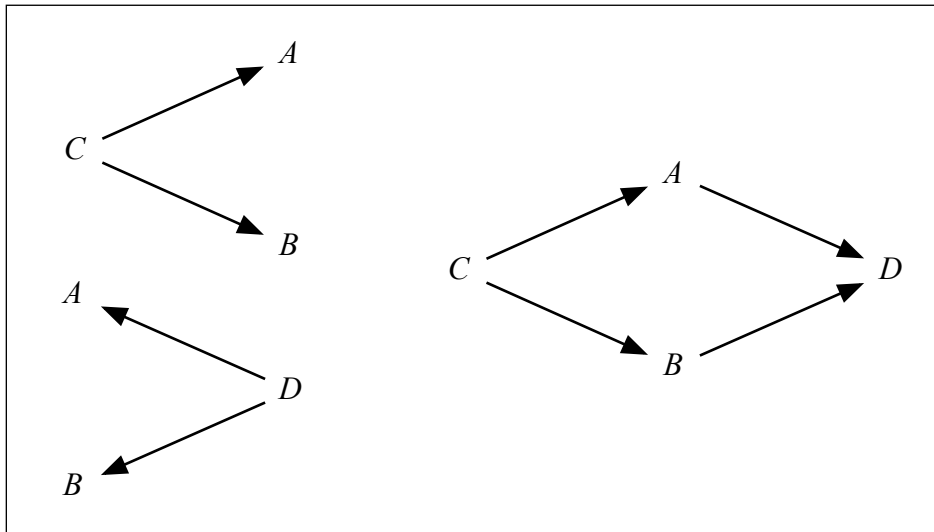
Figure 1: Open and closed forks.

cidences of events—our geyser-enthusiast recording, in another column, the times when the moon is full and the geysers spout, say—will satisfy the screening off relations (5)–(8)[7]. But this doesn't count as a conjunctive fork in Reichenbach's sense, since presumably the frequency of geyser spouts during a full moon is just what we would expect given the individual frequencies of geyser spouts and full moons generally. The fact that it is only probabilistically correlated events that stand in need of causal explanation will become important in what follows.

The existence of closed forks provides one example of a set of events that satisfy the probabilistic formulation of the principle of the common cause, and yet cannot form the basis for causal asymmetry—namely, the set of events in a closed fork comprising the correlated events, and the future screening off event. In this case the appeal to a past screening off event allows the principle to survive without appeal to temporally asymmetric facts. However there are other cases where the probabilistic relationships obtain and yet we do not have a conjunctive fork. One example is given by Salmon (1980, p. 217) and Arntzenius (1999, n. 2). Consider a case where an event C is the common cause of events A_1 and B_1, and where A_1 causes A_2 and B_1 causes B_2. In this case A_2 and B_2 will be screened off by both A_1 and B_1, and yet by stipulation neither is a common cause of their correlation (see Figure 2).

Likewise for causal chains (Papineau, 1993). The probabilistic relations defined above hold between any three items in a causal chain; when A causes C which causes B, or when B causes C which causes A, C screens off A and B (see

---

[7]Of course (5) and (6) follow trivially, since there was no correlation in need of screening off—the point here is that any instance of record keeping establishes the probabilistic relations (7) and (8).
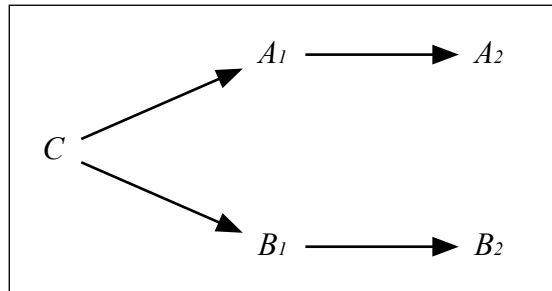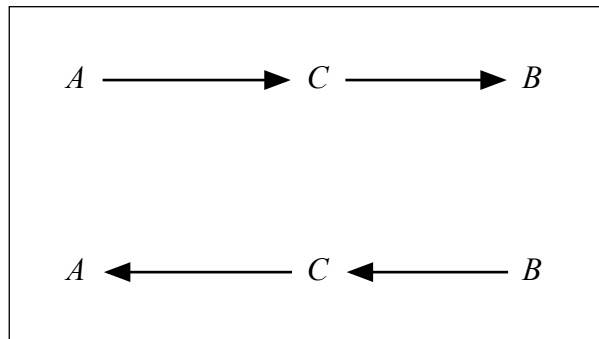
7

Figure 2: Screening off without a common cause.

Figure 3: Screening off in causal chains.

Figure 3). Indeed, Papineau (1989, p. 337) takes screening off of this sort to be in part definitive of the notion of a causal chain.

This is where the assumption of temporal orientation is standardly introduced. The assumption of temporal orientation can discriminate between the open conjunctive fork in Figure 1 and the causal chains in Figure 3, if we know the temporal order of the events in question, on the assumption that causes precede their effects in time. That is, if we take time to run from left to right, and assume that the conjunctive fork open to the future in Figure 1 correctly represents the temporal order of the events, then both chains in Figure 3 can be ruled out on the grounds that they represent a later event (*A* and *B* respectively) causing an earlier event (*C*). Likewise, if we assume that either chain in Figure 3 correctly represents the temporal order of the events, then Figure 1 can be ruled out on the grounds that it represents a later event (*C*) causing an earlier event (*A* and *B* respectively).

But of course we cannot use this method if it is the direction of causation itself that we are trying to ground probabilistically. There remains a weaker, symmetric temporal principle, which Reichenbach is sometimes held to have invoked, where *simultaneous* correlated events are ruled out as candidates for direct causal connections with each other (see for example Berkovitz, 2002, pp. 242*ff*; Hausman, 1998, p. 210 n. 2, also takes Papineau, 1989, p. 336, to be flirting with this idea).

The idea here is to appeal to special relativity, which forbids any direct causal connection between simultaneous events[8]—so if A and B are simultaneous then we can rule out both chains in Figure 3 on grounds that have invoked temporal order but not temporal orientation. This principle is far too weak to cover all the cases, however—often one or more of the effects of a common cause will occur earlier or later than the others, and intuitively the principle ought to apply to these as well. Indeed, the only earth-bound correlated events which would fall inside this constraint would be those occurring within the order of nanoseconds of each other (Salmon, 1980, pp. 217-218).

For those seeking a reduction of causation to probabilities, there is another solution available, and it invokes essentially the same strategy used for ordering closed forks as described above. The idea is to find some further event whose probabilistic relationship with the events in question enables the asymmetry to be determined. Take, again, Figure 1 and Figure 3, and suppose there is some further event (E) which satisfies the following probabilistic relationships (following Papineau, 1993, p. 240): correlated with C and B; not correlated with A; correlation with B screened off by C; correlation with C not screened off by anything. If we assume that correlation is essential for causation, we can identify the correct diagram as the first chain in Figure 3—if Figure 1 or the second chain in Figure 3 were correct we should see E correlated with A (see Figure 4).

This general strategy also traces to Reichenbach, and has recently been defended by Dowe (2000). And as suggested in Section I, while Spirtes *et al.* (2000) do not see themselves as involved in such a project, both Papineau (1993) and Field (2003) have taken their causal modelling framework to provide a means of drawing the causal arrows onto networks of probabilistic dependence in this way. In the following section, I examine the prospects for this strategy.

III. WHERE IN THE WORLD IS CAUSATION?

In order to establish the causal order for any particular set of events, for whom the probabilistic relationships underdetermine the appropriate causal model, an appeal is made to the surrounding causal network. Thus, the asymmetry of any particular causal relation is extrinsically determined. This is not without particular advantages. For a start, those who wish to rest the asymmetry on something like the common cause principle, and yet keep causal asymmetry intrinsic, are faced with immediate problems concerning instances of causation which do not appear to form the right sort of forks. While the common cause principle is plausible as a more or less intrinsic feature of situations with macroscopically correlated effects, such as stones dropping into ponds, the sort of forks needed for situations without such macroscopic effects do not seem capable of playing the role required of them. For

---

[8]I leave to one side here issues concerning the possibility of simultaneous causation—even if simultaneous causation is possible in some circumstances, it certainly isn't going to apply to all cases which fall under the principle of the common cause.
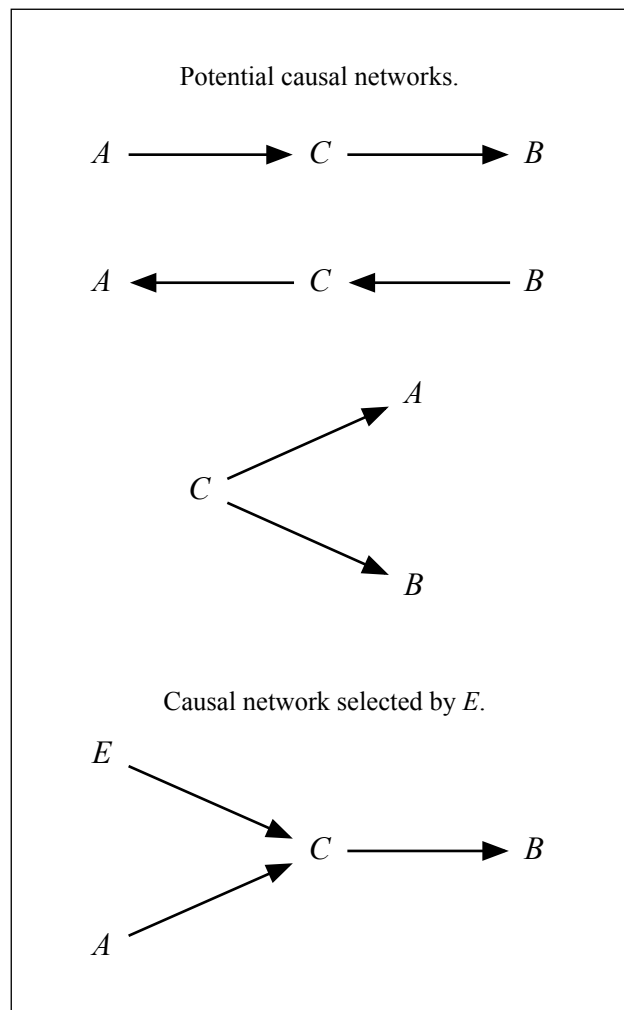
Figure 4: Fixing temporal orientation with further probabilities.

instance, take an inexpert billiard player shooting a cue ball into an eight ball on an otherwise empty table, and consider the causal relationship between the collision, and the eight ball hitting the cushion of the table. For Lewis (1979, 1986c), whose overdetermination is designed to provide the intrinsic asymmetry in every particular case of causation, it must be the case that the effects of the collision overdetermine the collision, while the causes do not. But where is this to be found in the region of the collision and its consequences? It cannot be in the macroscopic variables such as momentum, coarsely specified, and so Lewis ends up appealing to facts that in this case amount to the dissipative friction of the billiard balls on the table, and the light reflected from their surface (Lewis, 1979, pp. 469-470). These temporal asymmetries in the collision seem slender grounds for the causal asymmetry, however. If the collision occurs in the dark, at zero gravity, at zero

temperature, in a soundless vacuum, are we led to doubt the asymmetry of the case? Rather, it seems plausible to appeal to surrounding macroscopic events—the person hitting the ball, for example—and to trace out a network that will eventually exhibit the desired asymmetry.

It might be objected here that the move to the surrounding probabilistic network makes the asymmetry of causation *unacceptably* extrinsic[9]. But since it is already the existence of the second prong of a fork which gives the asymmetry to the first, and since it was a further extrinsic move which saved closed forks from being counterexamples to the principle of the common cause, this in itself will not be of concern to proponents of the approach. Moreover, if we adopt something like a frequentist view of probability, the theory is extrinsic right there in the probabilistic foundations (as it were). (On the other hand, if our theory of the causal relation is not itself a probabilistic one, we might have problems justifying why probabilities should matter for the asymmetry. This point is made by Dowe (1992b) with respect to the causal process theory of Salmon—the problem is why it should be that a causal process, which is an intrinsic property of a physical system, should be given its direction by extrinsic, *de facto* relations with surrounding causal processes. This point carries over to other theories which attempt to use the common cause principle as a plug-in solution for causal asymmetry, but needn't concern us here). Finally, as Papineau (1989, p. 336) suggests, we are already familiar with the extrinsic explanation of asymmetries, since the thermodynamic and radiative temporal asymmetries are plausibly explained by extrinsic boundary conditions[10].

A more specific concern about the appeal to extrinsic facts in fixing the direction of causation is that it may threaten the possibility of backwards causation—and since the theories under consideration are motivated in large part by the desire to

---

[9]Tooley(1987, p. 237; 1993, p. 22) has pressed the objection that to rely on causal nets makes causation unacceptably extrinsic. Appeal to intuitions concerning the intrinsic nature of causation has also been made by, for example, McDermott (1999, p. 303) and Lewis (1986b, pp. 205-207). A somewhat related concern is raised by Price (1993). Price points out that in order for temporal asymmetry not to be smuggled into the account (what he calls disguised conventionalism), the probabilities used must be temporally symmetric. But, he claims, if we use a naïve actual frequentist interpretation of probability, we become committed to only talking about causation where we have enough correlation to speak of statistical dependence—committed, that is, to the impossibility of single-case causation. It seems to me that this is a worry about probabilistic or regularity theories of causation in general rather than about their prospects for explaining causal asymmetry, however— Hume's account of constant conjunction is open to the same sort of objection, after all. The way out is, obviously, modal, and Price further charges that whatever modal notions are appealed to here will be as difficult to provide a temporal asymmetry for as the causal asymmetry we are seeking to ground. But this has no purchase on the theories under consideration, since the probabilities involved are all atemporal— if they were not, we would have been cut short at the very first step in the proposed reduction.

[10]While this is correct, the analogy shouldn't be pressed too hard. While we have a clear idea of what it would take for thermodynamic and radiative asymmetries to be reversed, it is less clear what criteria we should use to adjudicate cases of backwards causation. So for the former asymmetries, we have clearly defined asymmetric phenomena which stand in need of explanation; while in the latter case I take it that we are still attempting to explain the sense in which the phenomena is asymmetric in the first place.

11

make the existence of backwards causation an *a posteriori* matter, this ought to be particularly worrying. For example, there have been some suggestions for fixing causal asymmetry by indexing it to the temporal asymmetry of the neutral kaon, or the global entropy gradient provided by the second law of thermodynamics[11]. Given that these strategies fix the direction of causation globally, however, there can be no room for local instances of backwards causation—indeed, given this fact, it is hard to see how these strategies are an improvement over simply identifying causal order with temporal order. But the appeal to the probabilistic network does not have this particular problem. As we have seen, the probabilistic relationships definitive of open forks are defined atemporally, and so it remains a possibility that there exist situations with forks open to the past, even in the context of a network where the majority of forks are open to the future. So rather than ruling out the possibility of backwards causation, the theories under consideration provide the means of identifying when we have it—namely, anywhere there is a fork open to the past. Unfortunately, as I will argue in Section IV, this consequence of the approach is open to counterexample.

Before setting out the counterexample, however, there is a further issue for the causal modelling approach, related to extrinsic concerns, which bears examination. We can think of the issues discussed so far in this section as concerning where to place causal asymmetry. Is it an intrinsic property of causal relations, or is it determined extrinsically? So far, we have been concerned with how widely we need to cast the net. It turns out, unsurprisingly, that in order to carry out the reduction, we need nothing less than every relevant variable in the whole universe (Papineau, 1993)—or at least all those in the causal network in which we find ourselves. This follows, trivially, from the fact that what we consider a direct causal relation might in fact be one governed by a common cause of which we haven't taken account; and the fact that, as we have seen, the selection of the correct model from a series of candidates is achieved by appeal to further variables. The appeal to ever-wider networks has been a point of criticism (see for example Hausman, 1998, pp. 219-221), but it seems to me that the criticism relies on metaphysical intuitions of precisely the sort Papineau wishes to resist (see Section I). There are, however, issues not just regarding the *horizontal* placement of the asymmetry of causation, but also regarding what we can think of as the *vertical* placement of causation—concerning where the explanation of causal asymmetry lies on the scale from the microscopic to the macroscopic.

The worry here is raised by Price (1992), who notes that the fork asymmetry is absent at the microscopic level, given the time-symmetric determination of fundamental physics. It is for this reason that Field (2003) emphasises what he calls the *salience condition*:

> The salience condition needs emphasis: if the universe is two-way
> deterministic as in classical physics, one can find very unnatural vari-

---

[11]Dowe (1992a) suggests both strategies in the context of process theories, while Collins *et al.* (2004) suggest the entropic strategy for fixing counterfactual dependence.

> ables for which the temporal orientation [in causal graphs] is reversed [...] And with "exact" variables in the sense explained above [that is, variables specifying a complete physical description], the asymmetry completely disappears in classical physics.

Both Papineau (1993) and Field (2003)—exemplary in their commitment to *a posteriori* reduction—just bite the bullet here, and claim as an empirical discovery the fact that causation is much less widespread than we had thought, holding only at the level of those variables we find salient. I think a strategy endorsed by Dowe (1992a) is perhaps preferable here, where we instead have the microscopic asymmetry determined by the macroscopic asymmetry which it composes, or with which it is probabilistically connected. But in either case, the absence of the fork asymmetry at the microscopic level is kept at bay by taking causation to be a relation that only holds amongst the sort of macroscopic variables that we ordinarily take an interest in[12]. At this point we may begin to wonder how objective a reduction this is turning out to be, appealing as it does to the anthropocentric notion of salience. And there is a more serious problem lurking around the corner.

In order to introduce the problem, it bears mentioning the form in which the common cause principle has made it into the methods used in causal modelling. Spirtes *et al.* (2000), Pearl (2000) and Woodward (2003) adopt a restricted version of the common cause principle as an axiom in the procedure for inferring causal relations from statistical data, in the form of the *causal markov condition*. As Hausman and Woodward (1999, p. 524) observe, this can be usefully considered as the conjunction of two claims:

(1) If X and Y are probabilistically dependent, then either X causes Y or Y causes X or X and Y are effects of some common cause Z.

(2) Conditional on its direct causes, X is probabilistically independent of everything except its effects.

The causal markov condition is superior to the original formulation of the principle of the common cause in many respects, though we can safely set these aside for present purposes[13]. Here we can note firstly that the use of causal concepts in the characterisation makes it clear that it is not intended to be reductive, and secondly that it does nevertheless retain the claim that there is no probabilistic dependence without causal dependence (1), and that the effects of common causes are screened off by those common causes (2). A counterexample to (2) has been proposed by Salmon; my counterexample will be to (1).

Salmon (1980, p. 223) describes a case where our inexpert billiard player, in the same situation as described earlier, has only a half chance of sinking the eight ball. But suppose the case is such that for all the ways in which the eight ball might be sunk, the cue ball will certainly sink as well. Call the shooting of the cue ball C,

---

[12]This consequence of the approach ought to appear striking to those philosophers used to formulating causal exclusion arguments premised on causation being the province of fundamental physics.

[13]See Hausman and Woodward (1999) for the details.

the cue ball sinking A, and the eight ball sinking B. Here we have $P(B|A) > P(B)$, and therefore a correlation which stands in need of causal explanation. But the only *prima facie* candidate for a common cause—the shooting of the cue ball—doesn't screen off the correlation, since $P(B|C) = \frac{1}{2}$ while $P(B|A\&C) = 1$. Somewhat remarkably, given the line of argument for salience given above, the stock reply here, given by both Spirtes *et al.* (2000, p. 63) and by Hausman and Woodward (1999, pp. 528-529), is to appeal to a more complete specification of the physical system in question. Indeed, Hausman and Woodward explicitly *disavow* salience:

> The claim that is defended in the response is that [...] there must exist some set of screening-off common causes. [...] this is very different from saying that even the full set of variables that people ordinarily describe as common causes will screen off their joint effects or that it will be possible to specify a set of screening-off common causes in terms of any particular framework or vocabulary for dividing up the world [...].

It may be reason enough to suspect the general validity of the causal markov condition, that it doesn't hold among the variables we find salient (indeed, it was enough for Salmon to reject it). At the very least, it necessitates some substantial theoretical maneuvering on the epistemological side of the theory—in this example, how do we know that the cue ball sinking isn't a direct cause of the eight ball sinking, or vice versa, given that we don't have access to the precise physical details of the case? But the real problem here is that, as we have already seen, reasoning from the bottom up tells us that the causal markov condition will in fact cease to hold as we move to more fundamental, less inherently statistical levels of description. There is pressure here, then, from both above and below—the reformulation of the causal markov condition as a commitment to the existence of *some* set of variables for which it holds has all the flavour of a promissory note that cannot be cashed.

To reiterate the point here, we can look at Hausman and Woodward replying to another counterexample, given by Arntzenius (1993). Arntzenius observes that in a gas moving to thermodynamic equilibrium, there will be correlations between the temperature and pressure in different regions of the gas without there being a common cause. Again, Hausman and Woodward (1999, p. 530) revert to the "complicated causal story [...] involving huge numbers of molecules [...] that, because it is deterministic, must conform to the Markov Condition"—assuming determinism, they say, the "full set of determining causes" is such as to function as the required common cause. But notice that once we have zoomed down to this level, the full set of microscopic variables at *any* timeslice will suffice to perform this function—if we ignore temporal ordering, the probabilistic information radically underdetermines the choice of causal model, even if (in fact, precisely *because*) we can assume what Hausman and Woodward term *causal sufficiency* (the assumption that we have taken account of all the relevant variables)[14]. Given determinism,

---

[14]Hausman himself makes essentially this point when he points out that in the deterministic case,

that is, any time slice of the system will, on some model, fulfill the causal markov condition for any other time slice of the system (see Arntzenius, 1993, 1999). So what we have here is really, rather than the failure of the causal markov condition to apply to the case, a situation where the causal markov condition fails to pick out one causal model among many—and in a set where the direction of the causal relation can vary freely with respect to the direction of time.

IV. Yet One More EPR Counterexample

In the preceding section the focus was on the screening off property of common causes, captured by claim (2) of the Hausman and Woodward definition of the causal markov condition. In this section I give a potential counterexample to claim (1) of the condition, which was:

(1) If X and Y are probabilistically dependent, then either X causes Y or Y causes X or X and Y are effects of some common cause Z.

It is well known that quantum mechanics raises a number of difficulties for this claim, and in what follows I focus on the famous EPR cases (Einstein *et al.*, 1935), and in particular the constraints on their interpretation provided by Bell's Theorem (Bell, 1964)[15]. I will not rehearse the experimental arrangements or the theorem itself here, as both have received extensive treatment elsewhere. Rather, I will describe a consequence of the theorem for the principle of the common cause that to my knowledge has not yet been noticed.

Recall the geyser-enthusiast from Section II, who by recording correlated geyser spouts created future screening-off events; and recall Reichenbach's claim that this is not a case of backwards causation so long as there exists a corresponding screening-off event in the past. Now, suppose that our geyser-enthusiast becomes an EPR enthusiast, recording in her notebook the measurement outcomes from a series of EPR experiments. Here we have a set of future screening off events, as in the geyser case. But in the EPR case, Bell's Theorem rules out the existence of any past screening off event (so-called hidden variables) with which to close the fork, and so we have a *prima facie* case of forks open to the past. If such cases (for EPR correlations are recorded all the time) in fact provide forks open to the past, then we have a set of clear counterexamples to the principle of the common cause.

Note that this result is stronger than is usually claimed by those who see the EPR cases as refuting the principle. Normally the claim is simply that we have in these cases a set of correlations which are not (perhaps cannot) be screened off— and therefore a counterexample to the universality of the principle. If this were the only problem, however, it could be easily evaded by making the formulation

---

"the probability of y conditional on the direct causes of x will be the same as the probability of y conditional on x and all the direct causes of x" (1998, p. 215)

[15]I use the phrase for the time being to refer not only to the specific theorem first given by Bell, but to the family of theorems inspired by Bell that purport to prevent any local-realistic interpretation of quantum mechanics. Later in the paper I will focus on one particular theorem.

conditional: *if* there is a screening-off event, *then* the direction of causation is given by the principle of the common cause. After all, the principle of the common cause is a principle governing *causation*, and—*pace* Reichenbach—it need be no part of such a theory that all correlations permit causal explanation. This is the attitude taken, for example, by Papineau (1989), who—referring to the EPR correlations as "unscreenable-off" (p. 336)—writes: "I don't claim that all correlations are causal, just that if there is a screener-off it is the cause" (*ibid*). The EPR enthusiast shows that this is false, by providing an example of forks open to the past which are clearly not instances of backwards causation—if they were, we would be able to manipulate past measurement results by writing in notebooks.

At this point it may seem as if something must have gone wrong—if Bell's Theorem shows that there cannot be a common cause of the EPR measurement results, why does it not apply equally well to the records taken by the EPR enthusiast?[16] And yet surely the case we have imagined is possible; for one, the EPR correlations are perfectly analogous to the Geyser correlations from the perspective of the EPR enthusiast; and of course, EPR measurement results are recorded with unremarkable frequency by experimental physicists the world over. This apparent tension is resolved by considering more carefully what Bell's Theorem rules out, and what the EPR enthusiast must write down in her notebook in order for her records to screen off the measurement results. Since there is a family of theorems which fall under the Bell moniker, and what is ruled out differs from theorem to theorem, I focus here for definiteness on the version due independently to Wigner (1970) and Belinfante (1973).

As catalogued by van Fraassen (1982, p. 31), the premises of the Bell-Wigner argument are *causality*, *locality* and *autonomy*. Causality is simply the common cause requirement: that the hidden variables function as a common cause of the measurement results, so that the joint probabilities of measurement results for the two particles are just the product of the probabilities of the individual measurement results, conditional on the hidden variables. Locality and autonomy are restrictions on how the requirement of causality may be satisfied, and so strictly are premises that go beyond the principle of the common cause[17]. Locality dictates that the measurement result for a particular particle depends only on the hidden variables and the apparatus measuring that particular particle, or to put it conversely, measurement results for a particular particle are independent of measurements performed on the other particle. Autonomy is the requirement that the hidden variables are statistically independent of the type of measurements performed on either particle—that is, hidden variables do not influence the selection of which measurements to perform, and measurements do not influence the values of hidden variables. To-

---

[16]I owe this observation to Huw Price.

[17]This bears noting in this context since van Fraassen (1982, p. 32) takes a failure of autonomy to entail a failure of the principle of the common cause in general. The availability of backwards causation models of quantum mechanics in the context of common cause theories of causation shows this to be false—see for example Dowe (1997), clearly a coherent if in my view untenable interpretation. See Suárez (2004) for further discussion.

gether, the assumptions capture the intuitive EPR view according to which pre-existing elements of reality determine measurement results; we are free to choose what to measure; and our measurements do not have non-local or backwards-in-time effects. The view is often referred to as *local realism*, and the violation of the Bell-Wigner inequality by quantum mechanics shows that at least one of the cluster of commitments that make it up must be discarded.

With the premises made explicit, it can be seen immediately which one is available for the purposes of our hypothetical EPR enthusiast—namely, correlations between the measurements performed and her notebook records, amounting to a failure of autonomy. The reason why this has been unattractive as a premise to give up for the purpose of making sense of the EPR experiments is that if we are considering a set of hidden variables as a potential common cause prior to measurement, to give up autonomy means either having the hidden variables dictate the selection of which measurement to perform, in violation of free will (and conspiratorially thwarting randomised experimental arrangements), or alternately having the selection of measurements determine prior values of hidden variables by backward causation—options which have been largely ignored by interpreters of quantum mechanics[18]. But if we consider a future screening off event, such as a record noting the type of measurement and the outcome, we can unproblematically appeal to the second of these options, since like any other recording device, what we have is a quite mundane case of forward causation.

The defence of the principle of the common cause at this point turns on finding some past screening off event to close the fork, and I will make some brief comments on two options for doing so. One option is to take the EPR enthusiast to demonstrate the existence of a past common cause, and Bell's Theorem to show that such a past common cause must entail a violation of locality or autonomy—and therefore seek to develop a model consistent with these constraints. The upshot of the EPR enthusiast, however, is that such a model cannot simply consist in finding an event that bears to the measurement outcomes the probabilistic relationships definitive of a common cause—in our recording device, we already have such an event, so we must be looking for something *more* than just this correlational structure. For example, it is not sufficient for the development of a backwards causation model of the EPR correlations to merely identify some past event which is correlated both with the measurement settings and measurement outcomes—some additional property of this event must serve to differentiate it from the recording device. The suggestion by Dowe (1997) that a backwards causation model consistent with the principle of the common cause demands hidden variables set by forks open to the past (that is, not just by correlation with measurement settings) is one way this

---

[18]The most plausible development of the former option appeals to variable detector efficiency, first proposed by by Pearle (1970) and most fully developed in the so-called prism models of Fine (1982b,a). The latter option dates back to O. Costa de Beauregard and has been physically most highly developed by Cramer (1997), and philosophically most developed by Huw Price—see Price (1984, 1994, 1995, 1996b); Price (1994) is criticised by Dowe (1996), with a reply by Price (1996a). See also Dowe (1997).

can be achieved.

A second option is defended by Hausman and Woodward (1999, pp. 565-567). They propose that the two EPR measurements are not distinct events, on the basis that there is no means of independently manipulating the individual measurement results. They accept that this is controversial, but here rather than challenging this step of their argument I will concentrate on the implications. While it is a possibility that the measurements at some level form a single event, clearly the measurement outcomes as recorded macroscopically—computer readouts, dials moving, and so on—are all paradigmatically distinct events. Since these latter events are correlated, the common cause principle demands a screening off event; and as before, our EPR enthusiast can happily provide one. Since, as before, this is clearly not a case of backwards causation, what this entails is that Hausman and Woodward are committed to the measurement event being the common cause of the measurement records. While they claim that their model avoids "causal pathologies" (p. 267), it is unclear that a spatially distributed event of this sort playing the role of an instantaneous common cause (instantaneous since the earliest distinct events will be spacelike separated) does any better—in fact, it seems that this really amounts to a kind of violation of locality.

The central point to make, regardless of the merits of any of these options, is that it is *not* an option for those who subscribe to the principle of the common cause to rule out the EPR cases as not falling under the principle, or not being the kind of correlations that call for a causal explanation. To the extent it is attractive to think of quantum mechanics as not permitting causal explanation of any sort at all, this is not a position that can consistently be held together with the principle of the common cause. Whether this is a reason to give up the principle, or rather to seek the proper causal explanation, I leave to the commitments of the reader.

V. Concluding Remarks

The principle of the common cause needs to be placed in the world both horizontally, via extrinsic networks of probabilistic relations; and vertically, by locating at what scale the variables satisfying the principle are to be found. Moreover, it demands that a causal explanation be given for the EPR cases in quantum mechanics. I have shown that these are tough demands to meet, by identifying a tension threatening the possibility of vertical placement, and giving an EPR counterexample which showed that ignoring these cases is not an option.

While I did not raise problems with horizontal placement in this paper, it bears noting in conclusion that there are other lessons to be drawn here. One of the initial attractions of the principle of the common cause is that it promises to provide something like a local reduction of the direction of causation. As it turns out, those wishing for a probabilistic reduction of causation via the principle end up requiring a wealth of extrinsic relations, so that the causal asymmetry between any two variables turns out to depend on their relationship to many more variables. The upshot is that those seeking such an objective reduction for the asymmetry of cau-

sation should not be hostile to rival views which also propose extrinsic accounts. In particular, they should not be over-hostile to agency views of causal asymmetry, which can in this light be seen simply as a different form of extrinsic account of the direction of causation.

REFERENCES

Arntzenius, Frank. 1993. "The Common Cause Principle", in Hull *et al.* (1993), pp. 227–237. 14, 15

Arntzenius, Frank. 1999. "Reichenbach's Common Cause Principle", in Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, Winter 1999 Ed, 1999. Online: http://plato.stanford.edu/archives/win1999/entries/physics-Rpcc/. 7, 15

Belinfante, Frederik Jozef. 1973. *A Survey of Hidden Variable Theories*, Pergamon, Oxford, 1973. 16

Bell, John S. 1964. "On the Einstein-Podolsky-Rosen Paradox", in *Physics*, Vol. 1, 1964, pp. 195–200. 15

Berkovitz, Joseph. 2002. "On Causal Inference in Determinism and Indeterminism", in Harald Atmanspacher and Robert Bishop (Eds.), *Between Chance and Choice: Interdisciplinary Perspectives on Determinism*, Imprint Academic, Thorverton, 2002, pp. 237–278. Online: http://www.umbc.edu/philosophy/berkovitz/papers/determinism.pdf. 8

Collins, John; Hall, Ned; and Paul, L. A.. 2004. "Counterfactuals and Causation: History, Problems, and Prospects", in John Collins; Ned Hall; and L. A. Paul (Eds.), *Counterfactuals and Causation*, MIT Press, Cambridge MA, 2004, pp. 1–58. Online: http://mitpress.mit.edu/books/chapters/0262532565chap1.pdf. 12

Cramer, John G. 1997. "Quantum Nonlocality and the Possibility of Superluminal Effects", in Marc G Millis and Gary Scott Williamson (Eds.), *NASA Breakthrough Propulsion Physics Workshop Proceedings*, 1997. Online: http://www.npl.washington.edu/npl/int_rep/qm_nl.html. Proceedings of a conference at the NASA Lewis Research Center in Cleveland Ohio, 12-14 August 1997. 17

Dowe, Phil. 1992a. "Process Causality and Asymmetry", in *Erkenntnis*, Vol. 37, 1992, pp. 179–196. 12, 13

Dowe, Phil. 1992b. "Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory", in *Philosophy of Science*, Vol. 59, 1992, pp. 195–216. 11

Dowe, Phil. 1996. "Backwards Causation and the Direction of Causal Processes", in *Mind*, Vol. 105(418), April 1996, pp. 227–248. 17

Dowe, Phil. 1997. "A Defense of Backwards in Time Causation Models in Quantum Mechanics", in *Synthese*, Vol. 112(2), August 1997, pp. 233–246. 16, 17

Dowe, Phil. 2000. *Physical Causation*, Cambridge University Press, Cambridge, 2000. 3, 4, 9

Ehring, Douglas. 1982. "Causal Asymmetry", in *The Journal of Philosophy*, Vol. 79(12), December 1982, pp. 176–774. 3

Einstein, Albert; Podolsky, Boris; and Rosen, Nathan. 1935. "Can Quantum-Mechanical Description of Physical Reality be Considered Complete?", in *Physical Review*, Vol. 47, 1935, pp. 777–780. 15 May 1935. 5, 15

Field, Hartry. 2003. "Causation in a Physical World", in Michael Loux and Dean Zimmerman (Eds.), *Oxford Handbook of Metaphysics*, Oxford University Press, Oxford, 2003, pp. 435–460. 3, 9, 12, 13

Fine, Arthur. 1982a. "Antinomies of Entanglement: The Puzzling Case of the Tangled Statistics", in *The Journal of Philosophy*, Vol. 79(12), December 1982, pp. 733–747. 17

Fine, Arthur. 1982b. "Some Local Models for Correlation Experiments", in *Synthese*, Vol. 50(2), February 1982, pp. 279–294. 17

Good, Irving John. 1961a. "A Causal Calculus (I)", in *British Journal for the Philosophy of Science*, Vol. 11(44), February 1961, pp. 305–318. 3

Good, Irving John. 1961b. "A Causal Calculus (II)", in *British Journal for the Philosophy of Science*, Vol. 12(45), May 1961, pp. 43–51. 3

Hausman, Daniel M. 1998. *Causal Asymmetries*, Cambridge University Press, Cambridge, 1998. 3, 6, 8, 12, 15

Hausman, Daniel M and Woodward, James. 1999. "Independence, Invariance, and the Causal Markov Condition", in *British Journal for the Philosophy of Science*, Vol. 50(4), December 1999, pp. 521–583. 13, 14, 18

Hull, David; Forbes, Micky; and Okruhlik, Kathleen (Eds.). 1993. *Proceedings of the Biennial Meetings of the Philosophy of Science Association 1992*. Philosophy of Science Association, East Lansing MI, 1993. 19, 21

Kistler, Max. 1999. *Causalit et Lois de la Nature*, Vrin, Paris, 1999. 4

Lewis, David. 1979. "Counterfactual Dependence and Time's Arrow", in *Noûs*, Vol. 13(4), November 1979, pp. 455–476. 3, 10

Lewis, David. 1986a. *Philosophical Papers*, Vol. II, Oxford University Press, New York, 1986a. 21

Lewis, David. 1986b. "Postscripts to "Causation"", in Lewis (1986a), pp. 172–213. 11

Lewis, David. 1986c. "Postscripts to "Counterfactual Dependence and Time's Arrow"", in Lewis (1986a), pp. 52–66. 3, 10

McDermott, Michael. 1999. "Counterfactuals and Access Points", in *Mind*, Vol. 108(430), April 1999, pp. 291–334. 11

Menzies, Peter. 1996. "Probabilistic Causation and the Pre-emption Problem", in *Mind*, Vol. 105(417), January 1996, pp. 85–118. 4

Papineau, David. 1985a. "Causal Asymmetry", in *British Journal for the Philosophy of Science*, Vol. 36(3), September 1985, pp. 273–289. 3

Papineau, David. 1985b. "Probabilities and Causes", in *Journal of Philosophy*, Vol. 82(2), February 1985, pp. 57–73. 3

Papineau, David. 1988. "Response to Ehring's 'Papineau on Causal Asymmetry'", in *British Journal for the Philosophy of Science*, Vol. 39(4), December 1988, pp. 521–525. 4

Papineau, David. 1989. "Pure, Mixed, and Spurious Probabilities and Their Significance for a Reductionist Theory of Causation", in Philip Kitcher and Wesley Salmon (Eds.), *Scientific Explanation*, University of Minnesota Press, Minneapolis, 1989, *Minnesota Studies in the Philosophy of Science*, Vol. 13, pp. 307–348. 8, 11, 16

Papineau, David. 1993. "Can We Reduce Causal Direction to Probabilities?", in Hull *et al.* (1993), pp. 238–252. 3, 7, 9, 12, 13

Pearl, Judea. 2000. *Causality*, Cambridge University Press, Cambridge, 2000. 3, 13

Pearle, Philip M.. 1970. "Hidden-Variable Example Based upon Data Rejection", in *Physical Review D*, Vol. 2(8), 1970, pp. 1418–1425. 15 October 1970. 17

Price, Huw. 1984. "The Philosophy And Physics Of Affecting The Past", in *Synthese*, Vol. 61, December 1984, pp. 299–324. 17

Price, Huw. 1992. "Agency and Causal Asymmetry", in *Mind*, Vol. 101(403), July 1992, pp. 501–520. 12

Price, Huw. 1993. "The Direction of Causation: Ramsey's Ultimate Contingency", in Hull *et al.* (1993), pp. 253–267. 11

Price, Huw. 1994. "A Neglected Route to Realism About Quantum Mechanics", in *Mind*, Vol. 103(411), July 1994, pp. 303–336. 17

Price, Huw. 1995. "Locality, Independence and the Pro-Liberty Bell". 10th International Symposium of Logic, Methodology and Philosophy of Science, Florence, 1995. http://arxiv.org/abs/quant-ph/9602020. 17

Price, Huw. 1996a. "Backward Causation and the Direction of Causal Processes: Reply to Dowe", in *Mind*, Vol. 105(419), July 1996, pp. 467–474. 17

Price, Huw. 1996b. *Time's Arrow and Archimedes' Point*, Oxford University Press, Oxford, 1996b. 17

Reichenbach, Hans. 1956. *The Direction of Time*, University of California Press, Berkeley, 1956. 2, 3, 5

Salmon, Wesley C. 1980. "Probabilistic Causality", in *Causality and Explanation*, 1998 Ed, Oxford University Press, Oxford, 1980, pp. 208–232. Originally published in Pacific Philosophical Quarterly, Vol. 61, January-April 1980, pp. 50-74. 7, 9, 13

Spirtes, Peter; Glymour, Clark; and Scheines, Richard. 2000. *Causation, Prediction and Search*, MIT Press, Cambridge MA, 2000, 2nd Ed. 3, 9, 13, 14

Steel, Daniel. 2004. "Empirical Analysis and Causal Pluralism; or, What in the World is Causal Structure?" Online: http://www.msu.edu/user/steel/causality.pdf. 4

Suárez, Mauricio. 2004. "Causal Processes And Propensities In Quantum Mechanics", in *Theoria*, Vol. 19(3), September 2004, pp. 271–300. 16

Suppes, Patrick. 1970. *A Probabilistic Theory of Causality*, North-Holland, Amsterdam, 1970. 3

Tooley, Michael. 1987. *Causation: A Realist Approach*, Clarendon Press, Oxford, 1987. 11

Tooley, Michael. 1990. "Causation: Reductionism versus Realism", in *Philosophy and Phenomenological Research*, Vol. 50(Supplement), Autumn 1990, pp. 215–236. 4

Tooley, Michael. 1993. "Introduction", in Ernest Sosa and Michael Tooley (Eds.), *Causation*, Oxford University Press, Oxford, 1993, pp. 1–32. 11

van Fraassen, Bas. 1982. "The Charybdis of Realism: Epistemological Implications of Bell's Inequality", in *Synthese*, Vol. 52, 1982, pp. 25–38. 16

Wigner, Eugene P. 1970. "On Hidden Variables And Quantum Mechanical Probabilities", in *American Journal of Physics*, Vol. 38(8), 1970, pp. 1005–1009. 16

Woodward, James. 2003. *Making Things Happen: A Theory of Causal Explanation*, Oxford University Press, New York, 2003. 13