NANCY CARTWRIGHT AND MARTIN JONES

# HOW TO HUNT QUANTUM CAUSES[1]

## 0. INTRODUCTION

In Bohm's version of the Einstein–Podolsky–Rosen (E.P.R.) paradox, a source prepares paired spin-1/2 particles in the singlet state. This is a state in which the total spin must be zero. Yet, when a measurement is made along any given direction, each of the particles must yield a value of $+1/2$ or $-1/2$. Hence when measurements are made along the same direction on both members of a singlet pair, even though the two measurements can be carried out at quite distant places, the outcomes must be strictly anti-correlated. For measurements along two different directions, quantum mechanics also predicts a correlation, though it is not total:

$$\mathrm{P}(a = \pm 1/2 \;\&\; b = \pm 1/2) = 1/2 \sin^2(\theta/2)$$
$$\mathrm{P}(a = \pm 1/2 \;\&\; b = \pm 1/2) = 1/2 - 1/2 \sin^2(\theta/2)$$

where $a$ represents outcomes of spin measurements on the left-hand particle along a given direction, $b$ are outcomes on the right-hand particle along some possibly different direction, and $\theta$ is the angle between the two directions. The famous inequalities of J. S. Bell are commonly taken to rule out the one obvious account of these distant correlations – that they are due to the action of a common cause. The justification for this claim involves Reichenbach's conjunctive fork condition. Bell shows that – consistent with the predictions of quantum mechanics, which have been confirmed in experimental tests – it is not possible to posit a "hidden variable," $\lambda$, which produces a factorization between the two detection events:

$$\mathrm{P}(a = \Sigma_a \;\&\; b = \Sigma_b/\lambda) = \mathrm{P}(a = \Sigma_a/\lambda) \times \mathrm{P}(b = \Sigma_b/\lambda)$$

It is questionable whether Reichenbach's criterion is appropriately applied in this case (Cartwright (1989), Cartwright and Chang, forthcoming). Nevertheless, doubts about the possibility of a common cause explanation have led to a serious consideration of the possibility that

*Erkenntnis* **35**: 205–231, 1991.

the correlations are due to some direct causal action between the two separated detection events, and this despite the fact that the two measurements can be simultaneous (i.e., space-like separated) and also that there are apparently no appropriate forces exerted between the two, no propagating mass, no transfer of energy – in short, no known way in which one could influence the other.

What kind of evidence could help us decide whether there is some peculiar kind of causal interaction between the outcome events in an E.P.R. setup? This is the topic we want to explore in this paper. Again the work of Reichenbach immediately comes to mind – both his probabilistic criteria for determining causal connections and the closely-related mark method. But we do not propose to take on the extremely ambitious project of investigating how Reichenbach's methods could in general be put to use to solve the E.P.R. causality paradox. Rather, we will look in detail at one specific proposal with which Reichenbach's ideas have been associated. The proposal – called the robustness test – was put forward by Michael Redhead (1987). We are interested in it for two reasons. First because the robustness criterion attacks the problem at the level we feel is essential at this stage: it provides a direct test for causality between the outcome events, in contrast to the more usual round-about reasoning that there must be such a connection because there can be no common cause, or alternatively, that such a connection is impossible because it would violate relativity. What is more, thinking about how robustness works leads to some important cautionary lessons about causal testing in general, and about the application of Reichenbach's methods in particular.

## 1. ROBUSTNESS

To begin, let us simply quote Redhead's formulation of the robustness condition. Redhead begins by laying out a dilemma we apparently face when we try to account for the E.P.R. correlations causally; we must seemingly choose between an explanation's involving only a common cause and one involving a direct stochastic causal link between outcomes. However, he interjects,

> this conclusion can be questioned if proper account is taken of the necessary conditions for a direct causal link. By robustness of a causal relation we mean the following: A stochastic causal connection between two physical magnitudes *a* and *b* pertaining to two separated systems *A* and *B* is said to be *robust* if and only if there exists a class of

sufficiently small disturbances acting on $B(A)$ such that $b(a)$ screens off $a(b)$ from these disturbances.

Denoting the disturbance acting on $B$ by $d$, then the first part of this condition can be rendered formally as

$$\exists D(\forall d \in D(\text{Prob}(a = \Sigma_a/b = \Sigma_b \ \& \ d) = \text{Prob}(a = \Sigma_a/b = \Sigma_b))).$$

A similar condition can be written down for disturbance[s] acting on A. (1987, pp. 102–3).

There are questions both of interpretation and of justification here, questions which are necessarily intertwined. What are the "disturbances" Redhead has in mind? Why "sufficiently small"? And why should a direct stochastic causal connection satisfy this condition?

Redhead offers a number of clues in different places, and others have offered further suggestions. We shall examine some of these interpretive possibilities with an eye to two questions in particular: *Is* robustness a necessary condition on direct stochastic causal links? And if so, can the robustness test be applied conclusively to the correlations inherent in the quantum-mechanical singlet state?

Let us list the avenues which are open to us. The first clue comes immediately after the passage just quoted. Redhead continues:

The requirement of robustness as a necessary condition for a causal relation means that suitably small disturbances of either relata do not affect the causal relation. This is essentially the basis of the mark method for identifying causal processes. The processes propagate small disturbances (marks) in a local event-structure in accordance with the causal law at issue. (1987, p. 103).

So in the next section we shall look at Reichenbach's mark method and its possible links with robustness.

David Papineau (1989a, 1989b, 1990) has argued that robustness can be understood as a screening-off requirement on links in a causal chain. Redhead's own use of the terminology of screening-of, and his suggestions that robustness involves the notion of independence from the cause event's own causal history[2] both support this reading, so it forms the subject of Section 3.

Finally, Redhead has more recently formulated the robustness condition in terms of the preservation of functional form:

[W]e need to introduce a necessary condition for stochastic causality, that sufficiently small disturbances of the cause do not affect the functional form of the causal connection. We will refer to this as the robustness criterion for a causal connection. (1989, p. 148).[3]

This suggestion is taken up in Section 4. We shall have only limited success with it, but it does suggest a parallel approach in terms of treatment groups and control groups in the experimental search for causal links, and we explore this in Section 5. Before drawing our conclusions we concentrate, in Section 6, on the problems involved in applying even a reasonable version of the robustness condition to the specific case of paired electrons in the singlet state[4].

## 2. THE MARK METHOD

As we have seen, Redhead initially characterizes robustness as meaning that small disturbances of the cause do not affect the causal relation, and remarks that this is "essentially the basis of the mark method for identifying causal processes". A later discussion of Salmon's 'rotating searchlight' example also suggests links to the mark method (Redhead (1989), p. 150). Can we then use the mark method to justify the claim that robustness is a necessary condition on the existence of a stochastic causal link?

Perhaps the most plausible way involves a simple syllogism. If all causally connected pairs of events were capable of mark transmission (i.e. could pass the mark method test), and all pairs of events capable of mark transmission were correlated in accordance with the robustness condition, then it would follow that all causal connections would satisfy robustness. Simple enough; the problem with this syllogism, however, lies in defending its premises. We shall consider each in turn.

The claim that mark transmission is a necessary condition on causal connectedness is clearly too strong if it is taken to imply that *any* mark made on the cause must show up on the effect (unless 'mark' is understood in a way which makes the claim tautologous). Is the claim, then, that there must always be *some* mark which will be transmitted? No real justification of this is to be found; it would also seem quite difficult to falsify experimentally. Reichenbach says only that the claim is "confirmed by many experiences", but this is inadequate support if we are to apply the condition to quantum-mechanical situations which have broken the pattern of many of our previous experiences. Furthermore, the corresponding mark transmission test will be difficult to apply, partly for the reason that we will always be left wondering if we have just failed to find the right kind of mark. Perhaps so far we have had the misfortune to try repeatedly marks which "rub off" somewhere

along the causal chain of events. We also need to know how to recognise the time-evolved version of the mark when it makes its appearance at the effect event, and this will not always be a trivial matter.

Our scepticism about mark transmission as a necessary condition on causal connectedness may be heightened when we note that the typical examples used to illustrate the mark method involve pairs of events which are genidentical. In Salmon's terms,[5] causal processes transmit their own structure, so that a modification of that structure is propagated with the process: billiard balls carry chalk marks and train roofs carry snow – at least for a while (Reichenbach, (1956) p. 201 and (1928) p. 138, respectively). But the situation in EPR does not seem to be of this type. The cause and effect events, if such they are, are usually taken to involve different systems. Thus we should perhaps have even less faith in our ability to find a mark that might be transmitted, and that we might recognise in the effect, for in EPR the causal chain (if even *that* notion is applicable) involves interactions between different systems. The mark has to survive those interaction events, and in a recognisable form.

The defender of robustness can perhaps reply that some of these difficulties in application actually dissolve in the EPR case. Redhead's proof of the non-robustness of the singlet state correlations is not obviously supposed to be an application of the mark method. Nonetheless, if we regard the disturbances, $d$, as candidate marks, then one might claim that the generality of the proof (involving an "arbitrary perturbation" of the particle involved in the putative cause event (1987, p. 103)) corresponds to trying out all possible marks. And as the probabilities for outcomes on the distant particle are unchanged by any such perturbation (a point to which we will return in Section 5), we can arguably conclude that no mark turns up at the would-be effect, time-evolved or not.

However, there is another difficulty involved in applying the mark method, and it is one which proves far more intractable in the E.P.R. case. We have to be sure that in applying the mark we are not destroying the very causal connection we seek. But any perturbation of the kind Redhead considers destroys the singlet state of the two-particle system, and there is good reason to suppose that that state is somehow involved in any ability the right-hand outcome has to cause the left-hand outcome. The reason is that outside the singlet state we do not observe the strange correlations which set us wondering in the first

place. This is a point we shall elaborate on and defend towards the end of our discussion.

A similar consideration raises considerable difficulties for a defence of the second premise of our syllogism. Suppose that we have a mark which, without completely destroying $x$'s ability to cause $y$, interferes with the causal connection in such a way that marked $x$'s are less successful at producing $y$'s. Now, in the probabilistic case, the claim about mark transmission must be roughly the following: If $x$ and $y$ are cause and effect, then it is possible to mark $x$ in such a way that, *whenever $y$ follows*, $y$ is marked. Given this, it is clear we could have a pair of events which passes the mark method test, and yet which fails to satisfy robustness, since $y$ might turn up less often when $x$ is marked, and yet on its appearance always display a mark. Thus the claim that robustness is a necessary condition on passing the mark transmission test seems open to simple counterexamples. Consider, say, the ringing of migratory birds for identification. We know that the swallows which appear in Africa are the ones which left Europe, because all the swallows we find in Africa are ringed. This is perfectly compatible with the information that fewer swallows than usual are to be found arriving in Africa after a ringing campaign has been waged in Europe. Ringing simply prevents some of the more unfortunate birds from making the journey successfully.[6]

Interestingly, Reichenbach himself is open to exactly this objection. In his later work on the mark method he lays down an assumption which is very like robustness:

> Assumption $\beta$: If a mark is made in $b$, then
> either $p(a'/b') = p(a/b)$
> or $p(a/b') = p(a/b)$
> (1956, p., 201)[7]

Here $a$ and $b$ are events, and primes indicate the marked versions of these events. Reichenbach is assuming that $a$ and $b$ are causally connected, but is not assuming anything about which is cause and which effect. What he says about assumption $\beta$ is essentially that the second equation holds if $a$ is the cause and $b$ is the effect, or if the roles are reversed but the mark is one that "rubs off", whereas the former equation holds when the mark shows in $a$. This last claim sounds very like Redhead's robustness thesis. When two events pass the mark method test, the conditional probability of the effect on the cause is

unchanged by the marking process. The story of the swallows is just as much a problem for Reichenbach's assumption $\beta$ as it is for Redhead's robustness condition interpreted in terms of the mark method.

The intuitive connection between robustness (and perhaps assumption $\beta$) and the mark method might seem to be this: robustness must be satisfied before the mark method test can even be applied, for it must be possible to find a mark which does not interfere with the putative causal link in question. *A fortiori*, robustness is a necessary condition on passing the mark transmission test. The swallow story would seem to refute this suggestion: provided it only interferes to some (possibly large) extent, a mark which interferes *can* be used to apply the mark test. But suppose the intuitive argument were right. Suppose robustness were a necessary condition on even applying the mark method. We claim that this would only succeed in casting doubt on the first premise of the syllogism, that all causally connected events must be able to pass the mark transmission test. For by the second premise, we will not even be able to apply the test in the case of any non-robust causal connection there may be; we could hardly expect them to pass it. It is no use to protest that there are no non-robust causal connections, for that would only be to beg the question.

Overall, then, our best attempt to link robustness to causal connectedness via the mark method seems fraught with difficulties. There is little reason to think the one premise true, and good reason to think the other false. If the first premise were true, it is not clear that it would be of any use to us in the E.P.R. case. And if the second premise were true, it would seem only to endanger the first.

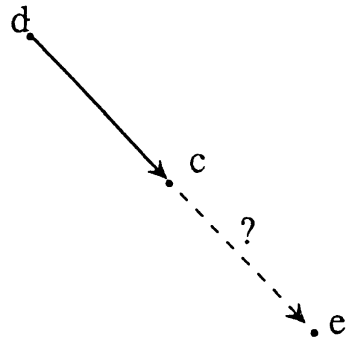
In passing, we should mention a different way of using the mark method to justify Redhead's claims about robustness. Passing the mark method test is more plausibly a sufficient than a necessary condition on causal connectedness – can we put it to use in that form instead? Unfortunately not. We could lay aside our previous worries and assert once more that robustness is a necessary condition on the mere application of the mark method, but this would only make robustness a necessary condition on a sufficient condition, and such a thing does not amount to a necessary condition. This way of connecting robustness to the mark method would simply mean that in non-robust cases we could not apply a test which, were it applicable, might reveal the existence of a direct stochastic causal link.

We conclude that the mark method cannot be used to justify the

claim that robustness is a necessary condition on the existence of a direct stochastic causal link. Perhaps more interestingly, however, we have had our first glimpse of a point that will recur, and prove a considerable obstacle to applying the robustness condition to E.P.R. on any interpretation. This point is that a test for causal connections can be impossible to apply unless we already know something about the causal details of our intervention, and even something about the causal structure of the situation under investigation.

## 3. SCREENING-OFF ALONG A CAUSAL CHAIN

David Papineau (1989a, 1989b, 1990) connects the robustness condition with a different aspect of Reichenbach's work. Papineau locates its significance in a familiar feature of Markov chains which Reichenbach explained: in a Markov process, intermediate causes screen off earlier causes from their later effects. We take it that the connection with Redhead's proof is made by casting the perturbation on the left-hand system as a (probabilistic) cause, $d$, of the measurement outcome, $c$, on that side, which is under consideration as a possible cause of the correlated measurement outcome, $e$, on the opposite wing. The causal picture looks like this




Papineau points out that the Redhead formula

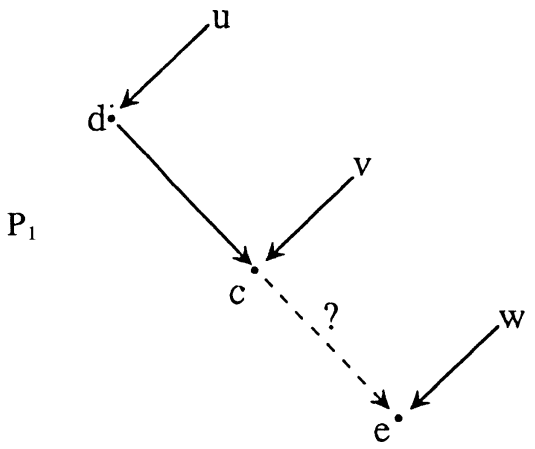
$$R : P(e/c \ \& \ d) = P(e/c)$$

is equivalent to the conventional Markov-inspired formula that requires that information about an initiating cause in a chain be irrelevant to the effect once a later cause in the chain is given.

Papineau's observation is provocative, but we think that it does not

help explain why robustness should count as a criterion of causality. It is true that the formula *R* holds when *c* is temporally intermediate between *d* and *e* in a particular constrained class of Markov models. But this seems to be irrelevant to the question of whether *c* is a cause of *e* in models in this class. Assuming that *d* (precedes and) causes *c*, and *c* precedes *e*, the models for which R holds satisfy two constraints. (1) They are Markov; (2) they contain no unbroken path from *e* to *d* *except* those that go through *c*. But this class contains both those models where there is a path that goes through *c* (i.e., *c* causes *e*) and ones in which there is not (i.e., *c* does not cause *e*). Satisfying R is thus irrelevant to whether *c* causes *e*.

The simplest structure in this class is probably also the one closest to the intended interpretation.




This picture is like the original, with the addition of letters from the end of the alphabet – *u*, *v*, and *w* – one for each of the already included factors, *d*, *c*, and *e*. These terms represent all the otherwise unmentioned causes of the factor in question and they thus provide information about the larger causal situation in which the small picture P is embedded. They are introduced subject to a convention. For any two factors which are unconnected, as *u* and *v* in $P_1$, it is assumed that the entire causal histories of these factors are independent – no unbroken route through the past of the later factor leads to the earlier, even allowing routes that go forward and backward in time.[8] In addition, a further assumption is made which underlies all this work connecting causal claims with probabilities – any factors (in any combination) whose histories do not intersect will be uncorrelated.

The idea behind the Markov assumption is that earlier causes always operate through intermediates, so that if you knew all the influences that bore on a particular effect at some time $t$, additional information about what happened prior to $t$ would be probabilistically irrelevant to the occurrence of the effect. In the context of the kind of causal diagrams that we are using here, the assumption can be formulated this way: If some node is held fixed on every causal path leading into $e$ (with the possible exception of the totally independent path from $w$, which may or may not have a node held fixed), then any information in the causal past of any of the fixed nodes will be probabilistically irrelevant to $e$.

Return now to diagram $P_1$. $P_1$ makes assumptions that are natural to Redhead's treatment of the E.P.R. setup. The perturbation, $d$, is taken as a cause of $c$, but there is no way that $d$ can cause $e$ except via the possible route from $c$ to $e$. In turn, $d$ is entirely produced by a "random factor" that operates independently of any other possible influences on $e$. In this case it is true that, given the Markov assumption, $c$ screens off $d$ from $e$. But that is independent of whether $c$ causes $e$ or not. Consider the two alternatives:

(i) C does cause $e$. Since, barring $w$, the only causal path into $e$ is via $c$, the Markov assumption ensures that $c$ screens off $e$ from $d$, which is in the causal past of $e$. Hence $P(e/d \,\&\, c) = P(e/c)$.

(ii) C does not cause $e$. Under the rules of the game, $e$ is probabilistically independent of $c$ since there is no unbroken path from $e$ to $c$. So, too, is the conjunct of $c$ and $d$ since there is no unbroken path from $e$ to either of them. Hence the screening-off condition holds, since $P(e/c) = P(e) = P(e/c \,\&\, d)$.

In both cases, then, condition R obtains under the Markov assumption; and in both cases the reason is the same. In structure $P_1$, there are no causes of $e$ other than $c$ and $d$. This is true more generally. Whether R obtains in a causal structure that embeds P depends not on whether $c$ causes $e$ or not, but rather on what other routes are available from $e$ to $d$. $P_2$ is a case in point. It embeds P, but R fails in it.

Before turning to $P_2$, it is worth looking to see what happens in $P_1$ when the Markov assumption is not made. Again consider the two alternatives:

(iii) C causes $e$: there is a direct causal path from $d$ to $e$; and thus the two will be correlated. If the Markov property is not assumed,

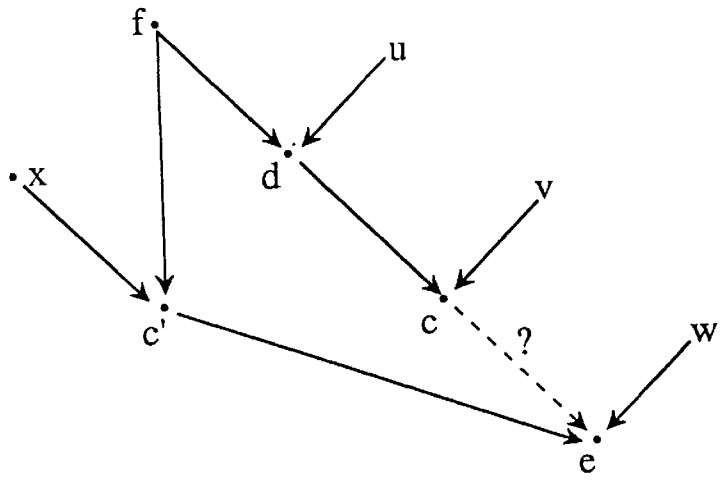
there is no guarantee that this correlation will be screened off by *c*. So R may fail.

(iv) C does not cause *e*: R obtains. The argument is the same as with case (ii), in which the Markov assumption played no role.

Surveying the four cases shows that the "causal intermediary" interpretation of robustness is not favorable to Redhead's intuition. For Redhead takes the failure of robustness as an indication that there is no causal connection. But in the situation pictured in $P_1$, the only possible case where R can fail is the third, and this is a case where the causal connection obtains.

Consider structure $P_2$ now, under the Markov assumption.

$P_2$




In $P_2$, both *c* and *d* may be probabilistically relevant to *e*, regardless of whether *c*, and thereby *d*, has a direct causal influence on *e* or not. Moreover, the relevance of *d* to *e* need not be screened off by *c*; this, too, is true regardless of whether *c* causes *e* or not. This is apparent from looking at the diagram. Imagine that *c* is known. Still *d* will be relevant to *e* since *d* is a symptom for the occurrence of *f*, which is itself a cause of *e*. The proof is tedious, but straightforward. Of course the structure $P_2$ is not a reasonable one to assume for Redhead's perturbation experiment, since, presumably, he would take *d* to be produced at the will of the experimenter, and to have no causes in common with *e*. We discuss it here merely to reinforce the point that condition R is in no way a criterion for the existence of a causal

connection between $c$ and $e$. Instead R holds or fails depending on what other factors cause $e$ and how these relate to $d$.

## 4. FUNCTIONAL FORM

As we have argued in Section 2, in order to use the mark method to test for a causal connection between two types of events, one already has to have a fairly detailed hypothesis about what the connection would be like if it did exist, and what kinds of marks it can propagate. The third suggestion for interpreting robustness – maintenance of functional form – requires less information about the putative connecting process, and hence provides a more generally applicable test. In the end we shall have to consider, of course, whether we do have the information we need for applying it to E.P.R. That will be the topic of Section 6.

Let us turn then to the question of functional forms. We begin by exploring the link with robustness in the simple case in which any causal connections present are deterministic. The setting of the account is important: we make sufficient observations to convince ourselves that two variables describing events in separate places are functionally related:

$$z = f(y)$$

We hypothesize that the relation between the two is due (entirely) to a direct causal connection between them.

> Hypothesis: $y$'s taking the value Y causes $z$ to take the value $f(\mathrm{Y})$.

A very simple paradigm for the causal connection is the rigid rod.[10] We can let $y$ represent the location of a red spot painted across the left end; $z$, the location of a blue spot on the right end. Imagine that a child is pushing the left end around in a circle. We, of course, are supposed to see neither the child nor the rod, but observe merely that the red spot executes the same circular motion as the blue. How do we test for a causal connection?

We follow standard experimental procedures,: *we* move the blue spot to a new position Y in the circle, a position that it would not occupy otherwise at that time due to the joint effects of all other causes operating on it; and then we look to see if the red spot moves in the

corresponding way as well. If it does, we are satisfied with our hypothesis; if not, we look for a more complicated causal story.

Three crucial assumptions must be made about our intervention if the procedure is to be a good test. Redhead leaves two unstated; the third, he acknowledges in his formulation of robustness. First, the intervention (call it $d$) must produce a value for the "cause" variable ($y$) that it would not otherwise have on that occasion. In the usual case we may not have information about what value $y$ would take otherwise. We try to make up for this by repeating the experiment a number of times, determining the type and timing of $d$ on each occasion in some "random" way. It is important to notice that these procedures do not constitute the test itself but are rather methodological tricks to help us achieve the ideal conditions of the test. To test the causal hypothesis we need to give $y$ a value it would not otherwise be caused to have, and it is always possible we may not obtain this result even though we have followed our instructions exactly. We might, for instance, choose our values for $y$ with some random number generator; nevertheless, by chance, it may pick just the values that $y$ would have had anyway for each run of the experiment.

The assumption that $d$ is "random" relative to $y$ and $z$ had a parallel in the "screening-off" approach to robustness, in that $u$, $v$ and $w$ were unconnected to one another in $P_1$. The second assumption we need to make was also broached there; it is simply that $d$ is incapable of affecting $z$ by any route other than via $y$. The failure of this assumption would be just as inimical to the successful application of this test as in the "screening-off" case.

Finally, the disturbance which we use to set the value of $y$ must not interfere with the (possible) causal process connecting $y$ and $z$. In our simple paradigm, that means that you have to keep the rigid rod *rigid*. If you move the red spot by bending the rod, you no longer have a test. The functional relation between $y$ and $z$ is not robust, but that does not show that there was no causal connection at work to begin with. Methodologically, we try to prevent destroying the causal connection by making the disturbance *small*, as Redhead indicates. Again, it is important to recall that smallness is not itself the requisite condition; it is just an indicator, and to no extent an altogether reliable one, that a necessary condition for the applicability of the criterion has been achieved.

The requirement for robustness of functional form can be taken as

a special case of Mill's methods of concomitant variations: we vary the cause and look to see if the effect varies in the concomitant way. But what counts as concomitant? That depends on the specific nature of the causal hypothesis under test. Our account began with an observed functional relation of the form $z = f(y)$, and proposed the most straightforward and limited explanation: $y$'s taking the value $Y$ causes $z$ to take the value $f(\mathrm{Y})$. Had we instead postulated a more detailed hypothesis of a specific connecting mechanism, like the rigid rod, we could have tested the new hypothesis not just by varying the cause across the range of its already observed value, but by trying out whole patterns of new behavior in the cause. In this case, the passively-observed functional form will be a special case, for a limited range of motions, of the function supposed to be governing the process; if the causal hypothesis is to be supported, we will demand robustness of the more general functional form. Note in particular that a direct causal link between $y$ and $z$ is compatible with the failure of the passively observed functional relation between them in the case in which some additional common cause, $x$, is operating: $z = g(x, y)$ might be robust whilst the *prima facie* relation $z = f(y)$ fails under experimental test.

What happens when we pass from deterministic causes to causes which are probabilistic? Standard causal modeling theory provides one very general account of how probabilistic causes might operate. The value of the effect variable (call it $x_e$) is represented as a deterministic function of the causal variable ($x_c$), but then a random "correction" factor ($u$) is added on. For example, we might have

$$x_e = mx_c^2 + u.$$

The distribution of the correction factor, $u$, then determines the distribution of the effect variable for a fixed value of the cause. For computational ease the correction term is usually taken to be normally distributed, with mean zero. In that case the value of the effect set by the deterministic function for a given value of the cause turns out to be the expectation of the effect for that value of the cause in the corresponding probabilistic case. In addition, it is usual to assume that the distribution of the correction factor is independent of the value of the cause. But this does not seem to be necessary to a concept of probabilistic causality. Changing the level of the cause could change not only the expectation of the effect, but the shape of the distribution around the expectation.

With respect to the possibility of a direct link between outcomes

in E.P.R., Redhead laconically tells us "In the present setup . . . the functional form of the causal relation is given by the conditional probabilities" (Redhead (1989), p. 148). We have already seen that the appropriate functional form for the application of the robustness criterion depends on the specific details of the causal hypothesis under test. What picture of causality must be assumed to settle on the conditional probabilities, as Redhead does? The standard causal modeling theory we have just been describing provides what might naturally be seen as an analogue of Redhead's criterion. It can be roughly expressed like this: the conditional expectation of $x_e$ given $x_c$ does not depend on the level of $x_d$, given that $x_d$ has no other route for influencing $x_e$ than via $x_c$. We must be careful though. Recall the discussion of Papineau's proposal for interpreting robustness using ideas akin to Reichenbach's time-ordering techniques. The condition we just stated holds whether or not $x_c$ causes $x_e$. The robustness of the conditional expectation of $x_e$ given $x_c$ under variations in $x_d$ shows not whether $x_c$ causes $x_e$ or not, but rather whether $x_d$ causes $x_e$ other than via $x_c$. When $x_d$ is an independent cause of $x_e$, robustness fails.

In the next section we will describe a way to think about robustness, via treatment/control tests, that does turn robustness into a necessary condition for a direct causal connection in certain cases, and which yields a closely-related criterion which is more generally applicable. It is based on the idea that, as you shift the probability of the cause, the probability of the effect should follow in train. An analogue of this result is derivable in conventional linear[11] causal models satisfying the conditions described above. Trivially, if

$$x_e = f(x_c) + g(x_0) + u$$

and

$$x_c = h(x_d) + v$$

then

$$\text{Exp}\, x_e = \text{Exp}\, f[h(x_d)] + \text{Exp}\, g(x_0)$$

where $x_0$ represents other causal factors, and $u$ and $v$ are random correction terms. When no other causal factors are at work, as Redhead supposes for E.P.R., then Exp $x_e$ = Exp $f[h(x_d)]$. This is an exact analogue of robustness as we interpret it in the next section: changes in Exp $x_d$ produce exactly the concomitant change in Exp $x_e$, so long

as $f(x_c)$ appears in the first equation. When it does not, Exp $x_d$ can vary as it will, with no effect on Exp $x_e$.

The significance of this result is uncertain though. It follows immediately on the decision to model stochastic events, like $x_e$, in these kinds of functional forms. Once we have done so, it is a trivial consequence that the probabilities for $x_e$ can be affected by interventions that change the probabilities for $x_c$ only so long as $x_c$ somehow gets into the equation for $x_e$. Thus we must consider two questions: (1) on what ground do we decide whether to admit laws with these functional forms into our theory?; and (2) why should we give them a causal interpretation? What the little result above shows, after all, is that the probabilities are robust if $x_e$ is a function of $x_c$ in our modeling equations. But this, on the surface, does not settle our fundamental question about whether nature has just set the relations which the equations describe between $x_e$ and $x_c$ as a brute fact, as Redhead in the end claims about E.P.R., or whether these relations are maintained by a causal relation between $x_c$ and $x_e$. To bear on this question, the equations need to be given some causal interpretation, and that is a very involved issue. Rather than detour through a long argument about causal modeling, we propose instead to think about robustness by appeal to more direct and immediate causal intuitions.[12] That is the project we turn to next.

## 5. TREATMENTS AND CONTROLS

There is one common way of using probabilities to test for causes that makes good sense of the robustness condition, and it is, moreover, probably the surest way to do so – that is the method which compares treatment and control groups. The method aims to solve problems of the type we face in E.P.R. A correlation between two factors is observed in a given population and the generalized version of the Reichenbach assumption is supposed: any true correlation must have a causal explanation. The idea of the treatment/control experiment is to establish the existence of a direct causal link between the two factors by eliminating all other possible causal accounts of their correlation. The strategy is familiar. We describe only the ideal version of it.[13]

In the control group, nature runs its course as usual. In the treatment group, the probability of the putative cause is altered artificially, in a way which is taken to be causally and probabilistically independent of any of the naturally occurring processes that are causally relevant either to the putative cause or to the putative effect. Label the experimentally

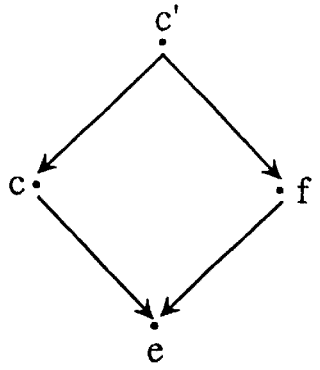
introduced factor which changes the probability of the putative cause, *d*; the putative cause, *c*; and the putative effect, *e*. In principle, then, *d* is supposed to have no other way of influencing *e* than via *c*; and *d*'s occurrence is assumed to be probabilistically independent of the operation of any possible joint causes for *c* and *e*, any joint preventatives, any preventatives of *e* singly, preventatives of *c*, causes of *c*, etc. In this case, when the number of *c*'s rises, so should the number of *e*'s, if *c* causes *e*; if *c* prevents *e*, the number of *e*'s should drop; and if *c* is causally irrelevant to *e*, the number of *e*'s should stay the same. So we expect

$$P(e) \text{ in treatment group} > P(e) \text{ in control group}$$

if *c* causes *e*.

To make a connection with the robustness condition we must think about the conditional probability, $P(e/c)$, in the two groups. To make any quantitative claims about the conditional probabilities, more assumptions are required. A standard assumption, often made without mention, is that the strength of *c*'s probabilistic propensity to produce *e* is the same whether *c* occurs naturally or is caused by *d*; all other factors, as well, are expected to function in the same way when *d* occurs as when it does not. Even with this condition, the robustness of the correlation between *c* and *e* will provide a criterion for their causal connectedness only in special cases. The special cases are those in which we can assume that the correlation between *c* and *e* is entirely due either to the fact that *c* causes *e*, or to the operation of other factors. The simplest kind of case this rules out is a causal structure like S, in which the correlation between *c* and *e* arises partly from *c*'s direct influence on *e*, and partly from the fact that *c* and *e* are joint effects of a common cause, *c'*.

S



In S, among the class of $c$'s, some of the $e$'s will be caused by $c$ alone, some by $c'$ via $f$, and some will be overdetermined, caused both by $c$ and by $c'$ via $f$. When we use $d$ to change the number of $c$'s, say to increase them, then if $c$ is a genuine cause of $e$, the number of $e$'s should increase as well. But the increase will not be proportional to their original ratio if $d$ and $c'$ are uncorrelated, as supposed. Under $d$, some of the $c$'s will be caused by $c'$ and these will still generate a component of the conditional probability, $P(e/c)$, due to the joint cause $c'$. But, some of the $c$'s will not be caused by $c'$, so the component due to the joint cause $c'$ will be missing in that class. All told, although the total number of $e$'s will increase as the number of $c$'s does, the conditional probability in a structure like S will fall. If instead $c'$ had been a cause of $c$ but a preventative of $f$, the conditional probability would rise. More complicated structures need to be considered, each on their own, but in any case what matters is how the values of $P(e)$ compare with and without $d$, not the values of $P(e/c)$. So in general, robustness is not a criterion of causality.

Here Reichenbach's thesis that conjunctive forks mark common causes plays a central role in the argument. We have already noted in the introduction that it is often supposed that there can be no common-cause at work in E.P.R. because the conjunctive fork condition fails. But we think this argument is mistaken: E.P.R. does not satisfy the background assumptions that must be met in order for the conjunctive fork to be a necessary condition for a common cause. So the question is left open as to whether a common cause exists, and hence it is an open question whether robustness should be expected even if the outcome in one wing of the E.P.R. setup is causing the outcome in the other wing. Robustness works as a test only under the "all or nothing" assumption; that is, under the assumption that if there is a direct causal connection between the two outcomes, this connection is entirely responsible for the correlation between the two, and no other factors contribute in any way.

Robustness aside, Redhead's proof now seems quite suggestive in itself. The basic idea of the treatment/control method is that changes in the probability of the cause should be associated with appropriate changes in the probability of the effect (given the conditions of an ideal experiment), and Redhead's proof bears as much on changes in the absolute probability of the effect as it does on robustness. It exploits the fact that the two systems in E.P.R. can be represented in different

Hilbert spaces. The fact that the intervention, $d$, is supposed to act only on the putative cause, $c$, and not directly on the effect, $e$, is captured by representing the intervention with an operator which acts on the space of the first system alone. It is a trivial fact built into the Hilbert space notation that operations on the one space have no effect on the other, even in situations where the two systems have previously interacted. In particular, there is no change in the probabilities on the second space.[14] That means that any changes in the probability of the cause will leave the probability of the effect intact. So it may be, after all, that standard treatment/control group methodology can rule out causation between outcomes in E.P.R., using the guiding idea of Redhead's proof. That depends of course on whether the conditions appropriate to the experimental test can be achieved. This is the question we turn to next.

## 6. CAN CONVENTIONAL CAUSALITY TESTS BE APPLIED IN E.P.R.?

In our discussion of functional form in Section 4, we pointed out three conditions that must be met before robustness could count as a test for causality. First, the cause must be changed so that it takes on a value it otherwise would not. The analogue to this condition in the treatment/control method is the requirement that the disturbance of the cause must be statistically independent of other causes, so that its action is not coincident with theirs. Second, the disturbance must have no way of influencing the putative effect other than via the candidate cause. This condition turned up in essentially the same form in the treatment/control test. Third, the change must be done in a way that does not distort the method by which the cause influences the effect. This is an essential feature, too, that must find its analogue in the conditions for a successful treatment/control test.

Consider a homely example. Sandy walks in the park every day. Whenever she leaves the park, a dog races after her, and leaves the park too. Is there a causal connection? Yes; and there is an enabling condition. Sandy and the dog have interacted in the past – it is her own dog, Kiwi, whom she has trained to respond to a high-frequency whistle that she carries along on her walks. We imagine a rough version of a treatment/control experiment: get Sandy to leave the park more often, and see if the dog does so as well. But we must be a little careful.

One easy way to get Sandy to leave the park an extra time is to get her to forget her whistle so that she has to go home to fetch it. In that case we will have succeeded in increasing the probability of the cause, but at the cost of destroying the mechanism by which it influences the effect.

The warning is equally apt in the quantum-mechanical case. One striking feature of the peculiar E.P.R. causality, if it exists at all, is that it exists only between systems which have previously interacted. It is assuredly not the case in general that getting a $+1/2$ outcome in a spin measurement on a particle in one location will cause a $-1/2$ outcome in a measurement on a particle in a distant location. The causal link seems only to operate between two particles which have interacted in the past, and which are now, as a two-body system, represented by a single quantum state, such as the singlet state of the E.P.R. arrangement. If the two systems can communicate causally, it is because there is something about their past interaction that enables them to do so.

How can we be sure that we are not interfering with any such enabling mechanism when we intervene to change the probability of the outcomes in one wing? Clearly we cannot be sure. We have not even the beginnings of a hypothesis about how the causal influence might be achieved, and the little we do know does not work in our favor. Any feature that enabled the two systems to influence each other would be a consequence of their past interaction. But we have no representation for any such unknown enabling factor. Our entire information about their past interaction is encoded in the quantum state – for E.P.R., the singlet state – and unfortunately, there is no way to change the probability of the outcomes on one system by operations on that system alone without destroying the very quantum state that records their past interaction. Redhead's intervention, $d$, rotates the spin states on the left and it thereby changes the probability of $+1/2$ in a spin-measurement along a given direction. But it simultaneously takes the pair out of the singlet state. What happens here is true in general, and for trivial reasons: you cannot keep the same state and yet get new probabilities. This is a point we first noted in the discussion of the mark method.

Perhaps this does not matter. Perhaps we can change the state that records the information about the past interaction without in any way perturbing specific features inherited from that interaction that might allow the two systems to influence one other. Perhaps not. We can

hope that a "small" intervention will not affect the candidate enabling feature, even though we know that no matter how small, it is bound to affect the quantum state. To tackle the problem we need some idea about what could make the influence possible. At the moment all we have is the quantum state, and it argues against the possibility of constructing a conventional probability test for causality.

Considerations about enabling factors may raise a further concern. The cause in our examples is, in one clear sense, underspecified. It is not just the occurrence of a +1/2 outcome that is supposed to be capable of influencing the outcome in a distant spin measurement. It is rather a +1/2 outcome plus an appropriate enabling feature. If we are serious about studying the possibility of a direct causal relation between the two outcomes, this observation may set us off in a new direction of investigation. But it should not make us any more sanguine about conventional testing for causality in our current state of knowledge. All we can propose now, by way of completeness, is that it is *a +1/2 outcome in a spin measurement of one member of a singlet state pair* that is the cause of a correlated outcome in the other member of the pair. Putting it this way again dashes our hopes for using the conventional test. Any action of the kind envisaged on one of the systems alone will change the quantum state. There is no way to increase the probability of *this* cause, so the test – though a good one when we can use it – cannot be applied.

Redhead, too, considers whether the quantum state should be included as a part of the cause or not, but he rejects the idea. He presents a dilemma: if the quantum state is causally involved in the production of a +1/2 outcome on the right, then either (1) it is a common cause of both the left-hand outcome and the right-hand outcome, which combines with a direct causal link between the two; or (2) it is a common cause of the two outcomes, and no other causes are at work. His argument against the second case relies on Reichenbach: the quantum state cannot be a common cause because it does not produce a conjunctive fork. We have already mentioned that we do not think this is a good argument, but do not want to take up that dispute here. In the end, the concerns we want to explore here bear as much on the first horn of Redhead's dilemma as on the second. So we turn to the first horn.

Redhead rules out this option by invoking the special theory of relativity: "... a direct causal link is unacceptable on relativistic

grounds" (1989, p. 149). But if the causal scheme represented in Redhead's first horn is to be rejected on grounds of conflict with relativity, then there would seem to be no need to apply the robustness condition in the first place. To the extent that we are concerned with the question of whether there is a direct causal link between the two outcomes, we might have settled the issue immediately by such considerations.

Turning the robustness condition from a necessary condition for causality, the status he usually claims for it, into a sufficient condition, Redhead also says about this horn: "If [the quantum state] is a partial common cause, then it must be overlaid by a direct causal link, since for fixed $|\psi\rangle$, Prob $|\psi\rangle$ $(a = \Sigma_a/b = \Sigma_b)$ has a robust form as a function of the values of $\Sigma_a$ and $\Sigma_b$" (1989, p. 149).[15] Presumably the robustness consists of the fact that for fixed $|\psi\rangle$, the conditional probability is not affected by outside disturbances since it is, after all, completely determined by $|\psi\rangle$ itself. But the same observation shows that robustness cannot be an indicator of causality in this case, since the interpretations that make it a valid test – preservation of functional form and treatment/control group testing – require that the disturbance affect the cause in the appropriate way; *then* we look to see if the connection between cause and effect is robust. But once $|\psi\rangle$ is fixed, we can no more affect either the occurrence of the cause or its probability by $d$ than we can affect the conditional probability.

This is surely a good thing, since, if it were a test, its results would be difficult to interpret. Recall from our discussion of structure S in Section 5, that where a common cause, like $|\psi\rangle$, and a direct cause, like the first outcome, are both at work, robustness is not to be expected. We may write, for this case:

$$\begin{aligned}P(e/c) = P(e\ \&\ c)/P(c) = [&P(c \text{ occurs } \&\ e \textit{ is caused by } c) + \\ &P(c \text{ occurs } \&\ e \textit{ is caused by } c') - \\ &P(c \text{ occurs } \&\ e \textit{ is caused by } c\ \&\ e \text{ is caused by } c')]/P(c)\end{aligned}$$

If the number of $c$'s is increased, the first term on the right, which is the conditional probability that $c$ causes $e$, should stay the same.[16] That means that $P(e/c)$ stays the same only if any changes in the last two terms cancel, which happens only when $c$'s causings of $e$ are totally coincident with those of $c'$. But that suggests a peculiar kind of collusion between the quantum state and the first outcome in their actions to bring about the second.

Redhead's overall conclusion from this reasoning is that the idea that the quantum state is part of the causal story in the production of the outcomes must be rejected. This frees him to use the lack of robustness of the conditional probability as a sign that the outcomes do not causally influence each other either, for he can now make the "all or nothing" assumption we described earlier. It seems, then, that no causes at all act in E.P.R. He says:

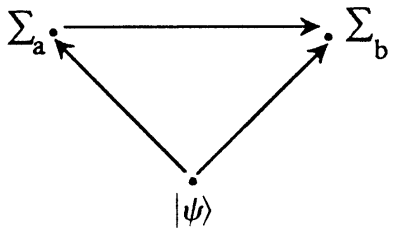
> So we have a dilemma, either horn of which is unacceptable. Our response must be to resist the temptation to restore robustness by regarding $|\psi\rangle$ as contributing a common cause, and accept the dependence on $|\psi\rangle$, not as a causal dependence, but as indicative of the nonrobustness of the link between $a$ and $b$, the way of passion rather than of action. (1989, p. 150)

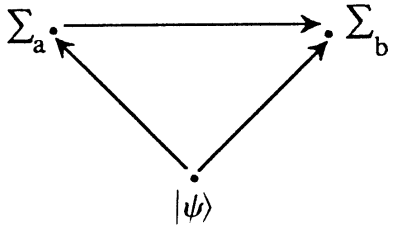
This reasoning, however, relies on a conflation of a simple and well-known distinction. Mill (1872, v. 2, bk. III, ch. x, s. 1) is the classical location for the distinction; it also lies at the heart of J. L. Mackie's work on *INUS* conditions (1980). Mill distinguishes between complete causes, which are separate and independent, and partial causes, which go to make up the different complete causes. Each of the complete causes is separately sufficient and together they are jointly necessary:

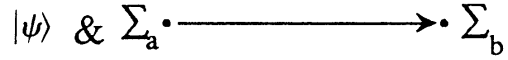
$$E \equiv C_1 \vee C_2 \vee \cdots C_n$$

Striking a dry match in the presence of oxygen is one way to get it to burn; lighting it from an already burning match is another; and so forth. Partial causes are the factors that go together to make up each complete cause – the striking, the dryness, and the oxygen, for example. These are Mackie's INUS conditions: *I*nsufficient but *N*on-redundant parts of an *U*nnecessary but *S*ufficient condition. Our complete causes, unlike Mackie's, are not logically sufficient for the effect, since they are supposed to operate purely probabilistically. But we still have the same distinction between complete probabilistic causes and the partial causes that make them up. Probabilistic causes have parts just as much as deterministic ones.

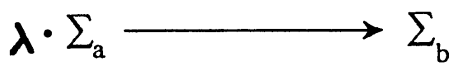
The upshot is that we have not two, but three, options if the quantum state is to play a role:

(1) 


(2) 


(3) 


Diagrams (1) and (2) correspond to Redhead's first and second horns, respectively. But it is diagram (3) that concerns us. We know, after all, that a left-hand spin-up outcome does not by itself affect spin outcomes at distant locations. Something more is needed, something which depends on the past interaction between the systems. Call it $\lambda$. So we know that if there is a causal link between the outcomes, a diagram like (3′) is needed:

(3′) 


But $\lambda$ is intimately connected with the quantum state, since the probabilities for how $\Sigma_a$ affects $\Sigma_b$ must be provided by it and not by the outcome $\Sigma_a$ itself. After all, these probabilities vary with the history of the particles for the same outcome value. In order to apply the rob-

hunt for quantum causes reminds us that his hopes were too optimistic. The mark method is not even a starter in testing for causal links between outcomes in E.P.R., because our background hypotheses about these links are too thin to supply the kind of information we need to put the method into play. When we turn to conventional statistical methods, we have seen that one test proposed – robustness of the conditional probabilities – can be conclusive only when we know that there are no other causal factors at work. In the particular case of E.P.R., it has often been assumed that this antecedent question can be settled by applying Reichenbach's conjunctive fork condition. But that application is in no way free of further modeling assumptions. Cartwright (1989) has shown that the conjunctive fork is only a necessary condition on a common cause under very limiting restrictions (restrictions that take one far from the case of maximal commonality); and she has argued that these special conditions are not satisfied in E.P.R.

Finally, even given the assumption that there are no other causes at work, the significance of robustness for E.P.R. is unclear. We need a model which tells us how one outcome would influence the other; without that, there is no way of interpreting the results of the robustness test so that it is decisive. In each of the approaches we discussed, Reichenbach provided the crucial guiding ideas that underlay our construction of a causality test; but the articulation of a specific criterion depends on the other details of the model. What is a criterion for a specific kind of link in one model need not be in another. We have illustrated with robustness and E.P.R., but we take the point to be perfectly general: there are no tests of causality outside of models which already have significant causal structure built in.

## NOTES

[1] We would like to thank Michael Redhead, Jeremy Butterfield, and J. B. Kennedy for several conversations about robustness, and David Papineau for the use of his manuscript.

[2] In conversation, Cambridge, 1990. See also the preface to the second (1989) impression of his (1987).

[3] This way of putting it is used again in the preface to the second impression of his (1987).

[4] In his (1989), Redhead provides two further clues, referring to David Lewis and Brian Skyrms. (1989, p. 148, n.2.) Space does not permit us to be thorough here; in our view, however, neither of these analogies holds. In the passage Redhead cites, Lewis is concerned with problems about individuating events – if we attribute too rich an essence

to them, certain causal counterfactuals will get the wrong truth value. So here there is just no obvious connection. In the case of Skyrm's notion of "resilience", the situation is arguably worse, for the verdict of that condition actually contradicts that of robustness in many cases.

[5] Salmon (1984), esp. Chapter 5. Redhead cites both Salmon and Reichenbach (1928 and 1956) in connection with the mark method. Redhead (1987), p. 118.

[6] This example has no pretensions to zoological accuracy.

[7] The notation has been changed in line with the rest of our notation.

[8] If their histories are causally linkable, the usual convention is to connect the two with a curved line, to which additional information may be attached. Most commonly this information includes the correlation between them. Without at least this information about the omitted factors, very little headway can be made in calculating what probabilistic relationships correspond to which causal relationships among the factors which are included.

[9] These are just the standard rules of causal modeling, which express just the kind of assumptions that Reichenbach himself made. We should point out that, without assumptions like this connecting causes and probabilities in the premises, it will not be possible to infer any such connections in the conclusions.

[10] This, in fact, is Redhead's own. Conversation, Cambridge, England, 1989.

[11] *Linear* means that the contributions due to each separate cause are additive.

[12] Also, causal modeling theory has the disadvantage that relations between "yes-no" quantities of the kind we see in E.P.R. are awkward to represent functionally.

[13] Real experiments will necessarily depart from the ideal since we lack either the knowledge or the requisite technique for much of what we would like to do in the ideal. In consequence real experiments, when well-designed, will be far more complicated than the ideal one since they need to control and correct for things we either do not know about or cannot manipulate as we wish.

[14] J. B. Kennedy has made essentially this point with regard to Redhead's "no-signalling" proof. Cf. Kennedy (1988) and Redhead (1987), p. 116.

[15] Note that if this is correct, then there can be no dilemma: the robustness Redhead mentions here would rule out the possibility that there is only a common cause, and no additional direct link between outcomes.

[16] As we point out in Section 5, this is an assumption that underlies all the reasoning here. Without it, robustness is not even a good starting idea.

## REFERENCES

Cartwright, N.: 1989, 'Quantum Causes: The Lesson of the Bell Inequalities', in *Philosophy of the Natural Sciences: Proceedings of the 13th International Wittgenstein Symposium*, Verlag Hölder–Pichter Tempsky, Vienna, pp. 120–27.

Cartwright, N. and Chang, H.: forthcoming, in a volume on realism edited by J. Margolis.

Kennedy, J. B.: 1988, 'Controllable Quantum Signalling', Lecture at the 13th International Wittgenstein Symposium, Kirchberg, Austria.

Mackie, J. L.: 1980, *Cement of the Universe*, Clarendon Press, Oxford.

Mill, J. S.: 1872, *A System of Logic*, Longmans, Green, Reader and Dyer, London.

Papineau, D., 1989a, 'Pure, Mixed, and Spurious Probabilities and Their Significance for

a Reductionist Theory of Causation', in P. Kitcher and W. C. Salmon (eds.), *Minnesota Studies in the Philosophy of Science, v.XIII: Scientific Explanation*, University of Minnesota Press, Minneapolis, pp. 307–48.

Papineau, D.: 1989b, 'Common Causes', M.S., Cambridge University.

Papineau, D.: 1990, 'Causes and Mixed Probabilities', International Studies in the Philosophy of Science, v. 4, no. 1, pp. 79–89.

Reichenbach, H.: 1928, *Philosophie der Raum–Zeit–Lehre*, Walter de Gruyter, Berlin, English translation, 1958, Dover, New York.

Reichenbach, H.: 1956, *The Direction of Time*, University of California Press, Berkeley and Los Angeles.

Redhead, M. L. G.: 1987, *Incompleteness, Nonlocality and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics*, Clarendon Press, Oxford, 2nd impression with new preface, 1989.

Redhead, M. L. G.: 1989, 'Nonfactorizability, Stochastic Causality, and Passion-At-A-Distance', in J. T. Cushing and E. McMullin (eds.), *Philosophical Consequences of Quantum Theory: Reflections on Bell's Theorem*, University of Notre Dame Press, Notre Dame, pp. 145–153.

Salmon, W. C.: 1984, *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, Princeton, NJ.

Department of Philosophy
Stanford University
Stanford, CA 94305
U.S.A.