**Relativity, Quantum Entanglement, Counterfactuals, and Causation**

Luke Glynn and Thomas Kroedel

(forthcoming in *British Journal for the Philosophy of Science*)

ABSTRACT

We investigate whether standard counterfactual analyses of causation (CACs) imply that the outcomes of space-like separated measurements on entangled particles are causally related. While it has sometimes been claimed that standard CACs imply such a causal relation, we argue that a careful examination of David Lewis's influential counterfactual semantics casts doubt upon this. We discuss ways in which Lewis's semantics and standard CACs might be extended to the case of space-like correlations.



## 1 Introduction

A source emits a pair of entangled particles, *A* and *B*, in opposite directions. Particle *A* travels to Detector 1, *B* to Detector 2. The detectors measure the particles' spins along the same axis, the possible values being 'up' and 'down'. The unconditional probability of *A* being measured 'up' is ½, as is the probability of its being measured 'down'. Likewise for *B*. There is a perfect (anti-)correlation between the spin values: the probability of *A*'s being spin up (down) *conditional* upon *B*'s being spin down (up), and vice versa, is 1. Experiments like this can be conducted in such a way that the measurement events are space-like separated. Variants on this experiment can be conducted in which *A* and *B* are measured for spin along different axes. Provided that the axes are non-perpendicular (a qualification that we'll leave implicit in what follows), the spin values remain (anti-)correlated, though not perfectly so.

Assuming orthodox quantum mechanics, the measurement results aren't 'screened off' (i.e., rendered probabilistically independent) from one another by the prior state of the particle pair together with the states of the two detectors. Since a strong case can be made that there can be no common cause (sufficient to explain the correlation) operating otherwise than via the prior state of the particle pair or the detector states, we are left with a correlation that isn't plausibly explained by common causes.[1]

Do standard counterfactual analyses of causation (CACs) imply that the measurement outcomes cause one another? While some have answered in the affirmative, we shall argue that a careful examination of Lewis's ([1979]) influential counterfactual semantics casts doubt upon this. We shall then discuss ways in which Lewis's semantics and standard CACs might be extended to yield clear-cut verdicts about such cases.

In more detail, the plan is as follows. Section 2 introduces the argument that standard CACs imply that the measurement outcomes cause one another. Section 3 argues that, when combined with Lewis's 'Analysis 1' (i.e., his 'asymmetry-by-fiat' analysis) of counterfactuals, CACs don't imply that the measurement outcomes cause one another, but rather fall silent on this matter. We take this to be a virtue, since it reflects the ambivalence of philosophers and physicists about applying our ordinary concept of cause to these cases. Section 4 argues that, when combined with Lewis's preferred 'Analysis 2' of counterfactuals (i.e., his 'closest worlds' analysis) plus his suggested similarity metric, standard CACs yield the decisive verdict that there is *no* causation between the measurement outcomes on the two entangled particles. Given our intuitive *in*decisiveness about whether such correlations are causal, we take this as a defect of Analysis 2. We therefore propose a retreat to Analysis 1. Finally, Section 5 shows that, while the combination of standard CACs with Analysis 1 succeeds in reproducing the indecisiveness of ordinary intuitions about the applicability of the concept of causation to these cases, the ordinary concept of causation might usefully be extended to yield determinate results about causation between the measurement outcomes. The idea would be to forge a revised concept of causation that might constitute a more useful part of the physicist's conceptual armory: in other words, to move from 'descriptive' analysis to 'prescriptive' or 'revisionary' analysis. Some possible revisionary analyses that take standard CACs as their point of departure are canvassed.

## 2 Measurement Outcomes and CACs

According to a plausible and highly influential tradition initiated by Lewis ([1973a]), causation is to be analyzed in terms of counterfactuals.[2] Analyses within this tradition typically take counterfactual dependence between distinct, actual events to suffice for causation.[3] (Event *e counterfactually depends* upon event *c* iff *e* wouldn't have occurred if *c* hadn't occurred.)

While plausibly sufficient, counterfactual dependence isn't necessary for causation. This is because of the possibility of cases of *redundant* causation, such as pre-emption and symmetric overdetermination (see Lewis [1986b], Postscript E), and of *probabilistic* causation where, in the absence of the cause, the effect would have had some residual probability of occurring (see Lewis [1986b], Postscript B). It seems clear that our case doesn't involve redundancy. We shall have more to say about probabilistic causation shortly.

Suppose that in fact particle *A* is measured spin up and particle *B* is measured spin down. Mainstream theories of events count *A*'s being spin up as a distinct event from *B*'s being spin down because they occupy distinct regions of space-time (Lewis [1986c]), or because *A* and *B* are distinct objects (Kim [1976]), since they occupy distinct parts of space when measured.

According to standard CACs, it is relevant to the question of whether the measurement outcomes cause one another whether *B*'s being spin down counterfactually depends upon *A*'s being spin up, and whether *A*'s being spin up counterfactually depends upon *B*'s being spin down. Whether these things are the case is a matter of whether (*CC*1) and (*CC*2) (respectively) are true:

(*CC*1) If *A* hadn't been spin up, then *B* wouldn't have been spin down.

(*CC*2) If *B* hadn't been spin down, then *A* wouldn't have been spin up.

Are (*CC*1) and (*CC*2) true? In the case where the particles are measured with respect to the same axis, Skyrms ([1984]) suggests that, in virtue of the law-like perfect (anti-)correlation between the measurement outcomes, both are true. He thus concludes:

> It appears that on [standard CACs] we must say that the measurement results […] caused each other, forming a rather odd, closed causal chain consisting of two spacelike-separated events. (p. 246)

Butterfield ([1992a]) argues that a natural generalization of standard CACs (proposed by Lewis ([1986b], Postscript B)) implies that there is causation between the measurement outcomes even in the version of the experiment in which the measurements are conducted with respect to different axes, where the (anti-)correlation is less than perfect. The natural generalization in question is designed to accommodate probabilistic causation. Instead of appealing to counterfactual dependence between events, this generalized analysis appeals to *probabilistic dependence*. Whether *B*'s being spin down probabilistically depends upon *A*'s being spin up and whether *A*'s being spin up probabilistically depends on *B*'s being spin down is a matter of whether (*CC*1*) and (*CC*2*) (respectively) are true:

> (*CC*1*) If *A* hadn't been spin up, then the chance (i.e., objective probability) of *B*'s being spin down would have been lower than it actually was.
>
> (*CC*2*) If *B* hadn't been spin down, then the chance of *A*'s being spin up would have been lower than it actually was.

Lewis (ibid. pp. 175-6) takes counterfactual dependence (captured by counterfactuals like (*CC*1) and (*CC*2)) simply to be a special case of probabilistic dependence (captured by counterfactuals like (*CC*1*) and (*CC*2*)). According to Lewis, (*CC*1) is true in the special case where, if *A* hadn't been spin up, the chance of *B*'s being spin down would have been zero.[4] Lewis (ibid., pp. 176–80) holds that probabilistic dependence in general suffices for causation.[5] In the absence of redundancy, it might also be taken as necessary. Butterfield argues in some detail—again by appealing to the lawful (anti-)correlations between the measurement outcomes—that, even where *A* and *B* are measured with respect to different axes, (*CC*1*) and (*CC*2*) come out true, committing Lewis to superluminal causation between the two measurement outcomes. Butterfield ([1992a], p. 28) takes this result to 'make trouble' for Lewis's analysis.

In what follows we shall focus upon the version of the experiment in which the particles are measured with respect to the same axis, and shall seek to cast doubt upon whether (*CC*1) and (*CC*2) come out true. Owing to the absence of redundancy and the deterministic nature of the connection in this case, standard CACs take counterfactual dependence to be necessary as well as sufficient for causation in this case, so doubt will thereby be cast upon whether standard CACs yield the result that the two measurement results cause one another. Though, to save space, we won't do so here, it is straightforward to generalize this reasoning to cast doubt upon whether (*CC*1*) and (*CC*2*) come out true in the case where the particles are measured with respect to different axes, and hence to cast doubt on whether Lewis's probabilistic generalization of standard CACs implies that the measurement outcomes cause one another in this alternative version of the experiment.

Before examining Lewis's semantics, however, we would like to point out that we are sympathetic to the positions of Skyrms and Butterfield at least to the extent that we agree that it wouldn't obviously be correct for a theory to entail that there is a causal relation between the measurement outcomes. There would, as Skyrms observes, be something 'odd' about such a causal relation. If Skyrms and Butterfield were correct in claiming that standard CACs entail such a causal relation, then the counterfactual analyst would have at least a *prima facie* difficulty in explaining why ordinary intuition seems to go indecisive concerning whether to interpret the relationship between the measurement outcomes causally.

The counterfactual analyst of causation might retort that she doesn't need to accommodate the indecisiveness of ordinary intuition about such cases within her theory. After all, the space-like correlations of quantum physics are far removed from our everyday experience, and so it is perhaps no surprise that ordinary intuitions are unclear about what to say about such cases.

A similar defense of his CAC was mounted by Lewis himself when discussing its failure to reproduce our alleged intuitive indecisiveness about whether symmetric overdeterminers count as causes. Lewis reasons as follows.

> When common sense delivers a firm and uncontroversial answer about a not-too-far-fetched case, theory had better agree. If an analysis of causation does not deliver the common-sense answer, that is bad trouble. But when common sense falls into indecision or controversy, or when it is reasonable to suspect that far-fetched cases are being judged by false analogy to commonplace ones, then theory may safely say what it likes. Such cases can be left as spoils to the

> victor, in D. M. Armstrong's phrase. We can reasonably accept as true whatever answer comes from the analysis that does best on the clearer cases. It would be still better, however, if theory itself went indecisive about the hard cases. If an analysis says that the answer for some hard case depends on underdescribed details, or on the resolution of some sort of vagueness, [Or on resolution of an ambiguity …] that would explain nicely why common sense comes out indecisive. (Lewis [1986b], p. 194; the text in square brackets is from Lewis's footnote that occurs at that point.)

Like the cases of overdetermination that Lewis was concerned with, the space-like correlations of quantum physics seem to be cases in which common sense falls into indecision or controversy.[6] Moreover, they are certainly not 'commonplace' in the sense of being objects of our everyday experience. So perhaps these count as 'spoils to the victor' cases: if CACs perform better than other analyses of causation in everyday cases, then we can reasonably accept their verdict about these more esoteric cases.

Still, Lewis notes that it would be *better* if theory itself went indecisive in cases where intuition goes indecisive. According to Skyrms and Butterfield, standard CACs don't go indecisive about space-like correlations between measurement outcomes, but yield the firm result that they are genuinely causal. In the next section, we will question this claim. Indeed, we will argue that, as desired, standard CACs do go indecisive about such cases, at least when combined with an appropriate semantics for counterfactuals. The indecision, we will argue, reflects an ambiguity in the counterfactual semantics in question.

## 3 Lewis's Analysis 1 of Counterfactuals ('Asymmetry-by-Fiat')

### 3.1 Analysis 1

In the previous section, it was observed that standard CACs take counterfactual dependence to be necessary and sufficient for causation in cases like ours, where there is no redundancy and where the connection between putative cause and effect isn't merely probabilistic. (Recall that the case on which we are focusing is the one where the measurements are conducted with respect to the same axis, so that the (anti-)correlation between the outcomes is perfect.)

This observation must be qualified, however, for counterfactual dependence is standardly taken to suffice for causation only if 'the right kind of counterfactual conditionals' (Lewis [2004], p. 78; cp. Lewis [1973a], pp. 565–7) are involved. The truth of *backtracking* counterfactuals—like 'If the barometer reading hadn't fallen, then the air pressure wouldn't have fallen earlier'—and *back-then-foretracking* counterfactuals—like 'If the barometer reading hadn't fallen, then (the air pressure wouldn't have fallen earlier and so) there wouldn't have been a storm'—doesn't suffice for causation. Consequently, according to standard CACs, the truth of (*CC*1) and (*CC*2) only implies that our space-like separated measurement outcomes cause one another if these counterfactuals express true foretrackers.[7]

Lewis ([1979]) describes an analysis of counterfactuals—his Analysis 1 ('asymmetry-by-fiat')—which is designed to yield the truth of only the foretracking counterfactuals appealed to by standard CACs.[8]

> *Analysis 1.* Consider a counterfactual 'If it were that *P*, then it would be that *Q*' where *P* is entirely about affairs in a stretch of time $\tau_P$. Consider all those possible worlds *w* such that:
>
> (i) *P* is true at *w*;
>
> (ii) *w* is exactly like our actual world at all times before $\tau_P$;
>
> (iii) *w* conforms to the actual laws of nature at all times after $\tau_P$; and
>
> (iv) during $\tau_P$, *w* differs no more from our actual world than it must to permit *P* to hold.
>
> The counterfactual is true iff *Q* holds at every such world *w*. (See ibid., p.

462)

By evaluating a counterfactual 'If it were that *P*, then it would be that *Q*' with respect to worlds that match the actual world (hereafter '@') in history up to $\tau_P$, we ensure that the counterfactual comes out false if it expresses a back- or back-then-foretracker. For example, a counterfactual 'If the barometer reading hadn't fallen, then …' is to be evaluated with respect to those worlds that match @ at all times before the time at which the barometer reading fell in @. In such worlds, the earlier fall in atmospheric pressure occurs (but the barometer malfunctions), and so the storm still occurs.

### 3.2 A frame-relative reading of Analysis 1

On a natural reading, some phrases that occur in Analysis 1, such as '*w* is exactly like our actual world at all times before $\tau_P$', make sense only relative to a given frame of reference. The result is that the verdicts of Analysis 1 depend on which reference frame one fixes upon.[9] When it comes to evaluating quotidian counterfactuals (like those concerning the barometer and the storm), nothing turns upon this ambiguity: these counterfactuals receive the same truth-values in all frames. But this isn't true of counterfactuals like (*CC*1) and (*CC*2) which concern space-like separated events.[10]

Condition (ii) of Analysis 1 tells us that the relevant worlds *w* to consider in evaluating (*CC*1) are those that exactly match @ in past history up until (but not including) the time at which *A* is measured spin up. On the relativistic picture, however, different frames yield different pasts. While the absolute past of *A*'s measurement is in its past relative to every frame, those events space-like separated from *A*'s measurement are in its past in some frames but not in others. What counts as past history, and what therefore is 'held fixed' in the relevant worlds *w*, affects what counterfactuals come out true. For example, let Frame 1 be a frame in which the measurements of *A* and *B* are simultaneous; let Frame 2 be a frame in which the measurement of *B* precedes that of *A*; and let Frame 3 be a frame in which the measurement of *B* occurs later than that of *A* (see Figure 1, where *a* and *b* are respectively the events of *A* and *B* being measured). Holding fixed the past of *A*'s measurement relative to Frame 2 requires holding fixed *B*'s being spin down. It immediately follows that counterfactual (*CC*1) comes out false on Analysis 1 if the relevant past to hold fixed is the past relative to Frame 2.

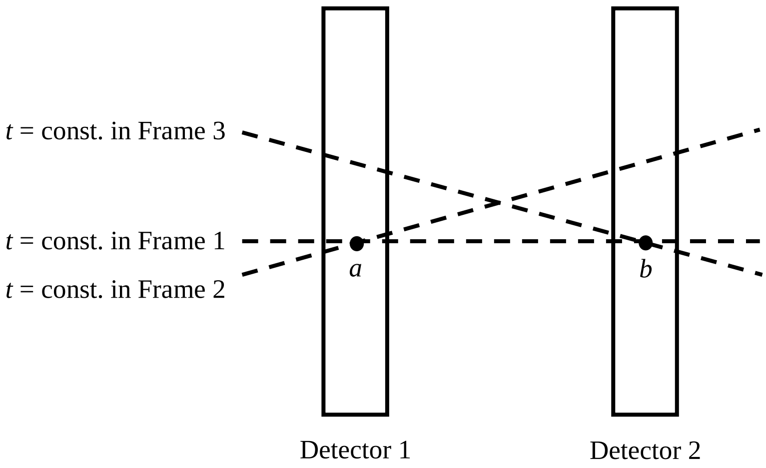


Figure 1

By contrast, holding fixed the past of *A*'s measurement relative to Frame 3 doesn't require holding fixed *B*'s being spin down. Indeed, counterfactual (*CC*1) comes out true if we hold fixed the past relative to Frame 3. To see this, note that, by condition (iv) of Analysis 1, the non-occurrence of *A*'s being spin up is to be brought about with minimal difference from @ at the time it is measured. Since *A* had chance ½ of being measured spin down, the minimally disruptive way of bringing about the non-occurrence of its being spin up is by having it be measured spin down.[11] By condition (iii), we are also to hold fixed the actual laws at all times after *A*'s measurement in this frame. This means holding fixed the lawful (anti-)correlation between the spins of *A* and *B*. In a world at which *A* is measured spin down and this (anti-)correlation obtains, *B* is measured spin up.

Counterfactual (*CC*1) also comes out true if the relevant past to hold fixed is the past relative to Frame 1. In this frame, *A*'s being measured spin up is simultaneous with *B*'s being measured spin down. Relative to Frame 1, past history up to (but not including) the time of *A*'s measurement again doesn't include *B*'s measurement. As was the case with Frame 3, *A*'s not being spin up can be achieved by its being measured spin down. The requirement of minimal difference (condition (iv) of Analysis 1) dictates that the lawful (anti-)correlation between the spins of *A* and *B* still obtain, so that *B* is measured spin up.[12]

By the symmetry of the case, (*CC*2) comes out false when evaluated with respect to Frame 3, and true when evaluated with respect to Frame 2 and Frame 1.

On a natural reading, there is thus an ambiguity in Analysis 1 when applied to space-like correlations, which results from the use of phrases such as 'all times before $\tau_P$'. Consequently, the truth-values of counterfactuals like (*CC*1) and (*CC*2) are indeterminate on Analysis 1. In virtue of this indeterminacy, standard CACs founded upon Analysis 1 go indecisive about whether space-like correlations are to be interpreted causally. This reflects the indecisiveness of intuition about these cases.

### 3.3 Frame-invariant readings of Analysis 1

Perhaps it is possible to resist the frame-relative construal of Analysis 1, so that its verdicts don't depend upon a choice of frame of reference. For example, one might read '$\tau_P$' so that it refers not to the *time* that *P* concerns but rather to the region of *space-time* that *P* concerns. For simplicity, assume that the antecedent of the counterfactual in question concerns a point-sized region of space-time.[13] Then that the relevant worlds are 'exactly like our actual world at all times before $\tau_P$' might be taken to mean *either* (i) that we are required to hold fixed the absolute past of $\tau_P$ *or* (ii) that we are required to hold fixed the complement of the absolute future of $\tau_P$, (excluding $\tau_P$ itself), that is, not just the absolute past of $\tau_P$ but also the absolute elsewhere of $\tau_P$.[14]

These two possible frame-invariant construals of Analysis 1 have contrasting implications for the truth-values of counterfactuals (*CC*1) and (*CC*2). To see this, suppose first that Analysis 1 is construed as instructing us to hold fixed the whole region of space-time outside the absolute future of *A*'s measurement. Since the measurement of *B* isn't in the absolute future of *A*'s measurement, the outcome of the measurement of *B* (namely *B*'s being spin down) is part of what gets held fixed. On this construal of Analysis 1, counterfactual (*CC*1) therefore comes out false. Similarly for (*CC*2), which comes out false as well.

Suppose that, by contrast, Analysis 1 is construed as instructing us to hold fixed merely the absolute past of *A*'s measurement. Since the measurement of *B* doesn't lie in the absolute past of *A*'s measurement, it isn't part of what gets held fixed, but instead is part of what is allowed to lawfully vary under the counterfactual supposition about *A*'s measurement. In virtue of the lawful (anti-)correlation between the spins of *A* and *B*, *B* is therefore spin-up in such a world. On this construal of Analysis 1, (*CC*1) therefore comes out true. Similarly for (*CC*2), which comes out true as well.

In general, if Analysis 1 is construed as instructing us to hold fixed the complement of the absolute future of the antecedent-event, then we don't get counterfactual dependence between space-like separated events. By contrast, if it is construed as instructing us to hold fixed merely the absolute past of the antecedent-event, we get counterfactual dependence (and indeed bi-directional counterfactual dependence) between the measurement outcomes on entangled but space-like separated particles.

### 3.4 Discussion

In the previous subsections, we have seen that Lewis's Analysis 1 of counterfactuals is multiply ambiguous when applied to a relativistic space-time. On one construal (which seems to us to be the most natural), the verdicts of Analysis 1 concerning the truth-values of counterfactuals like (*CC*1) and (*CC*2) become relative to a choice of reference-frame. Since Analysis 1 doesn't itself supply us with any grounds for making this choice (nor is there any obviously correct choice), Analysis 1, on this construal, is ambiguous about whether there is counterfactual dependence between our measurement outcomes.

But the ambiguity doesn't end there, for there is a further question of whether Analysis 1 is to be construed in this frame-relative way. There are frame-invariant readings as well. If Analysis 1 is to be construed in a frame-invariant way, then there is the still further question of *which* frame-invariant reading it is to be given. We outlined two possibilities: that the worlds that Analysis 1 takes to be relevant are those where we hold fixed merely the absolute past of the antecedent-event, or those where we hold fixed the complement of its absolute future.

The truth-values of counterfactuals which, like (*CC*1) and (*CC*2), concern correlated but space-like separated events depend upon the resolution of each of these open questions about how to interpret Analysis 1. As it stands, Analysis 1 therefore fails to yield determinate truth-values for these counterfactuals. In contrast, when it comes to evaluating everyday counterfactuals concerning time-like separated events, Analysis 1 yields the same verdicts no matter how these questions are resolved.

In cases like ours, standard CACs take counterfactual dependence to be both necessary and sufficient for causation. Since Analysis 1 fails to yield determinate truth-values for counterfactuals like (*CC*1) and (*CC*2), standard CACs that are founded upon Analysis 1 don't (*pace* Skyrms and Butterfield) yield the determinate verdict that the measurement events cause one another. Rather, they fail to yield a determinate verdict, thus reproducing the indecisiveness of intuition on this matter.

Indeed, not only do standard CACs that are founded upon Analysis 1 reproduce our intuitive indecisiveness about whether the outcomes of the measurement events are causally related, it seems that the indecisiveness of the resulting theory reflects an important reason for the indecisiveness of intuition, thus helping to satisfy Lewis's desideratum (quoted in Section 2) that the theory '*explain* […] why common sense comes out indecisive' (our italics).

Specifically, standard CACs take only foretracking counterfactuals to be relevant to causation, since—at least in ordinary cases—causes precede their effects. Analysis 1 is designed to yield the truth of only such foretrackers, and accomplishes this for ordinary cases by holding fixed history prior to the time of the putative cause event. But CACs founded upon Analysis 1 go indecisive about space-like correlations precisely because the time of the putative cause event (of *A*'s being spin up, say) is neither in the absolute past nor in the absolute future of the putative effect event (of *B*'s being spin down). It is consequently unclear whether the putative effect event is part of what is to be held fixed in evaluating the counterfactual (so that (*CC*1) comes out false), or whether it is to be allowed to vary lawfully with variations in the cause event (so that (*CC*1) comes out true). Different construals of Analysis 1 yield different verdicts.

The lack of a robust temporal priority of putative cause to putative effect—that is, the failure of the former to be prior to the latter in all reference-frames—seems to be a principal reason why physicists and philosophers typically hesitate to interpret the relationship between the outcomes of space-like separated measurements causally (see, e.g., Maudlin [2002], pp. 154–6).[15] CACs founded upon Analysis 1 of counterfactuals thus not only reproduce the intuitive indecisiveness about whether the correlated outcomes of space-like separated measurement events are causally related, but also go indecisive for the same reason (or at least for *a* principal reason) that intuition goes indecisive. Specifically, they do so for the reason that, unlike in paradigm cases of causation, there is no robust temporal asymmetry between the measurement outcomes. This suggests that, far from CACs running into trouble in the case of correlated outcomes of space-like separated measurement events, their treatment of these cases can reasonably be taken as one of their virtues.

## 4 Lewis's Analysis 2 of Counterfactuals ('Closest Worlds')

We have argued that standard CACs founded upon Lewis's Analysis 1 reproduce intuitive indecisiveness about whether the measurement outcomes should be regarded as causing one another. Nevertheless, Analysis 1 is in fact merely Lewis's first pass at giving the truth-conditions for counterfactuals. Lewis ([1979], p. 464) recognizes that it is built for a special case: counterfactuals whose antecedents concern localized events, evaluated with respect to ordinary, non-backtracking contexts. He consequently advances Analysis 2—his 'closest worlds' analysis—as a fully general analysis of counterfactuals.

> *Analysis 2*. A counterfactual 'If it were that *P*, then it would be that *Q*' is (non-vacuously) true if and only if some (accessible) world where both *P* and *Q* are true is more similar to our actual world, overall, than is any world where *P* is true but *Q* is false. (Ibid., p. 465, symbolism modified)

Lewis suggests that the similarity or closeness relation that combines with Analysis 2 to give the truth-conditions that counterfactuals receive in ordinary (non-backtracking) contexts is the one that is governed by 'weights or priorities' (1)–(4):

> (1) It is of the first importance to avoid big, widespread, diverse violations of law.
>
> (2) It is of the second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails.
>
> (3) It is of the third importance to avoid even small, localized, simple violations of law.
>
> (4) It is of little or no importance to secure approximate similarity of particular fact […]. (Ibid., p. 472)

Although Analysis 2 is more general than Analysis 1, Lewis claims that, when combined with the similarity metric governed by (1)–(4), it yields the same truth-values as Analysis 1 for the special case of counterfactuals with antecedents concerning localized events, which include those counterfactuals relevant to causation. But, as we shall now see, when we move to the relativistic case, the analyses aren't equivalent even in this special case. Unlike Analysis 1, Analysis 2 combined with priorities (1)–(4) doesn't make reference to temporal notions at all, but only to such things as laws and the size of regions of match. It consequently doesn't fall into the same indeterminacy as Analysis 1 if space-time is relativistic.

What results does Analysis 2 yield when combined with priorities (1)–(4) if space-time is relativistic? Consider a counterfactual 'If it were that *P*, then it would be that *Q*' where *P* is entirely about affairs at a space-time point $\tau_P$.[16] There are two obvious candidates for the closest type of world where *P* is true, which are similar to those we discussed in connection with the frame invariant construals of Analysis 1 in Subsection 3.3. Specifically, one candidate type is those worlds, $w_{\text{AP}}$ ('AP' for 'Absolute Past'), that exactly match @ in particular fact throughout the absolute past of $\tau_P$, but may differ from @ in particular fact in the absolute elsewhere of $\tau_P$ and in the absolute future of $\tau_P$. The other candidate type is those worlds, $w_{\text{AFC}}$ ('AFC' for 'Absolute Future Complement'), that exactly match @ in particular fact throughout the complement of $\tau_P$'s absolute future (though not in $\tau_P$ itself), but may differ from @ in particular fact in the absolute future of $\tau_P$.[17] Thus defined, all worlds that are of type $w_{\text{AFC}}$ are also of type $w_{\text{AP}}$, but not vice versa. (@ is trivially of both types.) Say that a world is *merely of type* $w_{\text{AP}}$ iff it is of type $w_{\text{AP}}$ without being of type $w_{\text{AFC}}$.

When comparing, according to (1)–(4), worlds of type $w_{\text{AFC}}$ with worlds that are merely of type $w_{\text{AP}}$, one thing that counts in favor of the relative closeness of worlds of type $w_{\text{AFC}}$ to @ is that they exactly match @ throughout a more extensive region of space-time than do the worlds that are merely of type $w_{\text{AP}}$. They therefore perform better according to priority (2), and will come out closer than worlds that are merely of type $w_{AP}$ provided that achieving this more extensive region of match doesn't require a big, widespread, diverse violation of actual law (a 'big miracle').

When it comes to counterfactuals like (*CC*1) and (*CC*2), at least a *small* miracle is required to secure exact match throughout the complement of the absolute future of the antecedent event. Consider (*CC*1). A world at which *A* isn't spin-up, but which exactly matches @ in the complement of the absolute future of *A*'s measurement is one at which a small miracle occurs to ensure that the space-like separated event of *B*'s being spin down still occurs, despite *A*'s not being spin up. Assuming this miracle is all that is required,[18] it seems that the closest worlds at which *A* isn't spin up, according to (1)–(4), will be of type $w_{AFC}$ and not merely of type $w_{AP}$ (cp. Bigaj [2006], pp. 93–6).[19] Counterfactual (*CC*1) therefore comes out false. The same points apply, *mutatis mutandis*, to (*CC*2), which comes out false as well.

There are, however, cases in which a single small miracle may not be enough to ensure exact match throughout the complement of the absolute future of *A*'s measurement. Suppose that *A* had been entangled for spin with *several* other particles, including *B*, which are measured in space-like separation from *A*'s measurement and from the measurements of one another. Then further miracles seem to be needed to break the lawful correlation of *A*'s spin with the spins of these other particles.[20] A large number of such miracles would seem to add up to a big miracle. The worlds of type $w_{AFC}$ where *A* isn't spin up thus involve a big miracle, since they match @ with respect to the measurement outcomes of the other particles, which occur in the absolute elsewhere of *A*'s being measured spin up. Worlds that are merely of type $w_{AP}$, by contrast, are allowed to differ from @ with respect to the measurement outcomes of the other particles. There are thus worlds that are merely of type $w_{AP}$ where the lawful correlation between the spins of all the particles is maintained and no extraneous violations of law occur. These worlds don't involve any big miracles and therefore appear to be the closest worlds where *A* isn't spin up according to (1)–(4). Analysis 2 consequently implies that the spin value of *B* counterfactually depends upon that of *A*. Given the symmetry of the case, Analysis 2 also implies that the spin value of *A* counterfactually depends upon that of *B*. In general, in cases where the closest antecedent-worlds are merely of type $w_{AP}$, we get counterfactual dependence (indeed bi-directional counterfactual dependence) between space-like separated measurement outcomes on entangled particles.

If, in the absence of redundancy and chancy connections, counterfactual dependence is taken to be necessary and sufficient for causation, the implications for causation are obvious: given Analysis 2 and priorities (1)–(4), in a case of simple bi-partite entanglement such as that originally described (where the closest antecedent-worlds are of type $w_{AFC}$), *A*'s being spin-up doesn't count as a cause of *B*'s being spin down. On the other hand, in a case of multi-particle entanglement (where the closest antecedent-worlds are merely of type $w_{AP}$), we get causal relations—indeed bi-directional causal relations—between the measurement results of *A* and *B*.

A CAC founded upon Analysis 2 thus doesn't do justice to our ambivalence about whether space-like quantum correlations are causal. On the contrary, it yields definite verdicts about such cases. It is interesting to note, however, that the definite verdicts that it yields are different from those that Skyrms and Butterfield take CACs to imply. In particular, we have argued that, in the two-particle entanglement cases upon which we have mainly been focused, the combination of standard CACs with Analysis 2 yields the verdict that that the spin values of the two particles are *not* causally related. So, if Analysis 2 is adopted then, *pace* Skyrms (*op*. *cit*.), we don't get 'a rather odd, closed causal chain consisting of two spacelike-separated events'. On the other hand, we do get odd bi-directional causal relations in multi-particle entanglement cases (which aren't discussed by Skyrms or Butterfield).

It isn't clear to us that the oddity of such relations is in itself a decisive objection to an analysis of causation that implies them.[21] But it is certainly puzzling that standard CACs founded upon Analysis 2 should yield these relations in certain multi-particle entanglement cases, but not in the case where only two particles are entangled.

Nevertheless, what we find most problematic about the combination of standard CACs with Analysis 2 is that the resulting theory yields clear verdicts about causation (including the verdict that there is no causation in the two-particle case) where intuition seems unclear. In virtue of their remoteness from ordinary experience, we might take these to be 'spoils to the victor' cases, about which theory may say what it likes. But certainly this is less satisfactory than having our theory reproduce our intuitive indecisiveness. In the previous section we argued that the combination of Analysis 1 with standard CACs does precisely this, and that the indecisiveness of the resulting theory is connected to a principal *reason* for our intuitive indecisiveness: that there fails to be a robust temporal asymmetry between putative cause and effect. This seems to be a strong point in favor of the combination of standard CACs with Analysis 1. We therefore take that combination to be superior to the combination of standard CACs with Analysis 2.

It might be argued that the combination of standard CACs with Analysis 2 sometimes does yield indeterminate results concerning space-like correlations. After all, the individuation conditions for miracles, and the number of small miracles required to constitute a big miracle, seem vague matters (see Woodward [2003], pp. 138–9). There may thus be cases—perhaps involving intermediate numbers of entangled particles—where it's indeterminate whether the closest antecedent-worlds for counterfactuals like (*CC*1) and (*CC*2) are of type $w_{\mathrm{AFC}}$ or merely of type $w_{\mathrm{AP}}$, resulting in a corresponding indeterminacy about the truth values of these counterfactuals and hence about whether there is causation between the spin-values of the entangled particles.

Our response is threefold. *First*, in extreme cases such as that involving just two entangled particles, the implications of Analysis 2 seem clear, even though intuition is indecisive about these cases too. Analysis 1, by contrast, is indecisive even about the two-particle case. *Second*, and relatedly, the reason why Analysis 2 goes indeterminate in cases involving intermediate numbers of entangled particles doesn't seem to have anything to do with the reason why our intuitions are indeterminate. It hardly seems that we are indecisive about whether to interpret space-like correlations causally because we are unsure whether it would take a large or a small miracle to alter the spin-value of one particle while holding fixed those of the particles with which it is entangled.[22] By contrast, as already noted, the reason Analysis 1 goes indecisive appears to be the same as a principal reason for which intuition goes indecisive (namely, the lack of a robust time-order). *Third*, and again relatedly, the vagueness of the distinction between big and small miracles, and hence the liability of Analysis 2 to go indecisive, has nothing particularly to do with space-like correlations in quantum mechanics. Analysis 2 is also liable to go indecisive for this reason in certain everyday cases involving time-like related macroscopic events where intuition delivers clear causal verdicts (see Woodward *op. cit*.). This provides an independent reason for preferring Analysis 1 to Analysis 2.[23]

## 5 Extending the Everyday Concept of Cause

Our concept of causation was shaped through experience of interactions between macroscopic events, where correlations can typically be explained in terms of causal relations between events that stand in time-like relations to one another. It should therefore come as no great surprise that intuition goes indecisive about whether the correlated spin properties of space-like separated particles cause one another. We have argued that the combination of standard CACs with Lewis's Analysis 1 does a good job of reproducing this indecision.

There is, however, a further question about the moral to be drawn from the indecisiveness of intuitions about such cases. Should we conclude that the concept of causation is simply not apt to fundamental physics, and that it ought to be dispelled from the inventory of concepts that the physicist draws upon in constructing her theories (though perhaps not from the conceptual inventory of the special sciences or of folk theory)?[24] Or should we maintain that the concept of causation has an important role to play in physics, but conclude that it must be revised and extended in the light of that physics? If the latter conclusion is drawn, the idea would be that we should seek to forge a revised concept of cause appropriate to what we now know about (for instance) the structure of space-time.

While we won't seek a definitive answer to these questions, it is worth noting that someone pursuing a revisionary (or 'prescriptive') CAC might achieve determinate results about space-like correlations by adopting one of the various precisifications of Analysis 1 that were discussed in Section 3.[25] Analysis 1 yields indeterminate results about counterfactuals like (*CC*1) and (*CC*2) because it specifies that the relevant worlds are 'exactly like our actual world at all times before $\tau_P$'. We saw that, in the context of a relativistic space-time, phrases like this are ambiguous. They might be construed as having any one of several precise meanings. In particular, that the relevant worlds are 'exactly like our actual world at all times before $\tau_P$' might be read as instructing us to hold fixed the absolute past of $\tau_P$, to hold fixed the complement of the absolute future of $\tau_P$, or to hold fixed the past of $\tau_P$ in a given frame of reference.

Combining a standard CAC with the stipulation that the correct worlds to consider in evaluating the relevant counterfactuals are those where we hold fixed the complement of the absolute future of the antecedent-event might be regarded as the most conservative extension of standard CACs to a relativistic space-time. It implies that our ambivalence about what to say about space-like correlations should be resolved in such a way that we don't admit them as causal. According to this option only pairs of events whose time-ordering is robust (as is the case for events that we regard as causes and effects in everyday life) are candidates for standing in a cause-effect relation.

Combining a standard CAC with the alternative stipulation that the correct worlds to consider in evaluating our counterfactuals are those where we hold fixed merely the absolute past of the antecedent-event might be regarded as a more radical option. In particular, as we have seen, this would be liable to generate bi-directional causal relations between the outcomes of our measurement events. Following Skyrms, one might take these bi-directional causal relations to be 'odd'. For our part, we don't regard the consequence of bi-directional causation to be a decisive objection against this strategy of holding fixed only the absolute past of the antecedent event. It is true that bi-directional causation isn't part of our everyday experience. But we are here concerned with possible extensions of the ordinary concept of cause to cases of quantum entanglement, which are quite foreign to ordinary experience.[26]

There is, however, a more interesting option for extending our ordinary notion of cause to space-like correlations in a relativistic space-time. In Subsection 3.2, we observed that the most natural reading of the phrases 'all times before $\tau_P$' and 'all times after $\tau_P$' that appear in Analysis 1 interprets these phrases as meaning 'all times before (after) $\tau_P$ relative to a given frame of reference'. On this construal, different choices of frames will lead Analysis 1 to yield different verdicts about counterfactual dependence between our measurement outcomes.

Suppose that we adopt this frame-relative construal of Analysis 1. There are then three natural options for extending standard CACs so that they yield definite verdicts about causation between our correlated but space-like separated measurement outcomes (which are subject to the usual qualification that the case doesn't involve redundancy or merely probabilistic connections):

(I) Counterfactual dependence in some frame is necessary and sufficient for causation.

(II) Counterfactual dependence in all frames is necessary and sufficient for causation.

(III) Causation is frame-relative. Counterfactual dependence in a frame is necessary and sufficient for causation in that frame.

The frame-relative construal of Analysis 1 combined with option (I) yields just the same results as the combination of standard CACs with the frame-invariant construal of Analysis 1 that stipulates the relevance of those worlds where we hold fixed merely the absolute past of the antecedent-event. That is, it gives us bi-directional causation between the spin-values of *A* and *B*.

The frame-relative construal of Analysis 1 combined with option (II) yields just the same results as the combination of standard CACs with the frame-invariant construal of Analysis 1 that stipulates the relevance of those worlds where we hold fixed the complement of the absolute future of the antecedent-event. That is, it gives us no causation between the spin-values of *A* and *B* (and indeed rules out causation between space-like separated events entirely).

The interestingly novel possibility is to combine the frame-relative construal of Analysis 1 with option (III). The resulting theory gives us not only frame-relative counterfactual dependence but also frame-relative causation between our measurement outcomes. Consider the three frames introduced in Subsection 3.2. The present theory implies that, in Frame 3, *A*'s being spin up causes *B*'s being spin down, but not vice versa; that, in Frame 2, *B*'s being spin down causes *A*'s being spin up, but not vice versa; and that, in Frame 1, there is bi-directional causation between *A*'s being spin up and *B*'s being spin down. One might find this option attractive if one thinks that the causal and temporal orders are tightly connected but also wants to do justice to the insights of relativity concerning the reversibility of the measurements' time-order.

Ultimately, however, our aim in this section has simply been to outline some possible alternatives for a counterfactual analyst of causation who wishes to extend the ordinary concept of causation so that it delivers determinate results about correlated events of space-like separation. A detailed evaluation of these various options must await another occasion.

## 6 Conclusion

We have examined the implications that standard CACs have concerning the space-like correlations of quantum mechanics when combined with Lewis's Analysis 1 and his Analysis 2 of counterfactuals. When combined with Analysis 1, standard CACs are indecisive about these cases, reproducing the intuitive indecision that many philosophers and physicists have felt about them. Indeed the reason that standard CACs go indecisive when combined with Analysis 1 corresponds to an important reason for our intuitive indecision, namely the fact that the measurement outcomes lack the sort of robust temporal asymmetry that is characteristic of paradigm cases of causation. This is a point in favor of the combination of standard CACs with Analysis 1.

That this combination reproduces our intuitive indecisiveness and does so for the right reasons is particularly notable given that it has previously been suggested (notably by Skyrms and Butterfield) that standard CACs run into trouble with space-like correlations, yielding firm verdicts of causation, despite the intuitive 'oddity' of these alleged causal relations. We have argued that, when standard CACs are combined with Analysis 1 of counterfactuals, this isn't so.

The combination of standard CACs with Lewis's Analysis 2 doesn't have the desirable property of reproducing the intuitive indecisiveness about space-like correlations (*a fortiori* it doesn't do so on the *grounds* for which intuition goes indecisive). On the contrary, in at least many such cases, it yields firm verdicts about causation. Interestingly, the firm verdict reached concerning the measurement outcomes on a pair of particles prepared in the spin singlet (viz., the verdict that there is no causation between their spin values) is the reverse of the verdict that Skyrms and Butterfield assume.

Finally, while the combination of standard CACs with Analysis 1 appears to fare well as a descriptive theory of causation (that is, it does well at capturing our intuitions about causation), there is the further question of whether we might seek to extend the everyday concept of causation to yield definite results about space-like correlations. Definite results can be achieved by combining an appropriate CAC with one of the precisifications of Analysis 1 surveyed in Section 3. We have remained neutral on which, if any, precisification (and choice of CAC) should be favored and on what the role of the concept of causation in physical theory should be.

## Acknowledgements

For helpful comments and suggestions, we would like to thank Frank Arntzenius, Adam Caulton, Christopher Hitchcock, Eleanor Knox, Dennis Lehmkuhl, Thomas Müller, Alastair Wilson, participants at a meeting of the 'Causality and Probability Research Colloquium' in Konstanz that took place in January 2011, at the 'Metaphysics Now' workshop that took place in Berlin in June 2012, and two anonymous referees. Luke Glynn's work on this article was supported by Deutsche Forschungsgemeinschaft grant SP279/15-1 and by the James S. McDonnell Foundation Causal Learning Collaborative.

Luke Glynn

*California Institute of Technology*

*Division of the Humanities and Social Sciences, MC 101-40*

*Pasadena, CA 91125*

*U.S.A.*

*lglynn@caltech.edu*

Thomas Kroedel

*Humboldt-Universität zu Berlin*

*Institut für Philosophie*

*Unter den Linden 6*

*10099 Berlin*

*Germany*

*thomas.kroedel@hu-berlin.de*

---

[1] See (Butterfield [1989], [1992b], [2007]; Hofer-Szabó [2011]) and references therein. The special case in which the particles are measured for spin along the same axis is in fact one to which hidden variable models positing a local common causal explanation are predictively adequate. Local hidden variable models don't, however, accurately predict the empirically observed correlations that arise where such particle pairs are measured for spin along different axes. It would be ad hoc to posit local hidden variables in the one sort of case, but not the other.

[2] Recent accounts in this tradition can be found in (Collins et al. [2004]).

[3] Hall ([2004]) distinguishes two 'kinds' or 'concepts' of causation, but argues that counterfactual dependence suffices for one kind.

[4] It is somewhat controversial to take, as Lewis (ibid.) does, (†) 'If it were that *X*, then the chance of *Y* would have been zero' as in general implying (‡) 'If it were that *X*, then *Y* wouldn't have occurred'. But the specific inference from 'If *A* hadn't been spin up, then the chance of *B*'s being spin down would have been zero' to (*CC*1) is plausible, since chance zero attaches to the particles having the same spin along the same axis in virtue of this outcome's nomic impossibility, and not because it is one of infinitely many possible outcomes.

[5] Strictly speaking, Lewis doesn't take mere probabilistic dependence to be sufficient for causation, but only what we might call *strong* probabilistic dependence: without the putative cause the probability of the effect would have been *much* less (i.e., less by a large factor) than

it actually was (ibid., pp. 176–7). Butterfield ([1992a], pp. 32–4) argues that this restriction shouldn't be regarded as saving Lewis's CAC from delivering problematic results concerning space-like correlations. It is worth noting that the claim that (strong) probabilistic dependence suffices for causation is more controversial than the corresponding claim about counterfactual dependence. For a survey of alleged examples of probabilistic dependence in the absence of causation, see (Hitchcock [2004]).

[6] At least this seems to be the reaction of common sense once a common-cause explanation is ruled out (as it is for reasons given in Section 1).

[7] Maudlin ([2002], Ch. 5) holds that counterfactuals like (*CC*1) and (*CC*2) are true of the sort of experiment we have described, but he doesn't attempt to discriminate back- and back-then-foretrackers from ordinary foretrackers (ibid., pp. 129, 129n). He consequently takes counterfactual dependence (of any of these stripes) to be sufficient not for causation, but for mere 'causal implication'.

[8] We present a version of Analysis 1 that is simplified in respects irrelevant to the discussion at hand.

[9] Finkelstein ([1999], pp. 290–1) briefly considers, but dismisses, a frame-relative interpretation of counterfactuals.

[10] In arguing that Lewis's CAC commits him to causation between the measurement outcomes, Butterfield ([1992a], esp. pp. 34–41, cp. [1992b], p. 51) assumes a fixed frame of reference, and doesn't discuss the implications that relativistic space-time has for the truth of the relevant counterfactuals. Neither does Skyrms (*op. cit.*).

[11] An alternative way is to have Detector 1 fail to register a reading. In that case, Analysis 1 still yields the truth of (*CC*1*) which, on Lewis's view, suffices for (probabilistic) causation.

[12] We assume that 'minimal difference' means minimizing violations of law even at the cost of sacrificing some minor match of particular fact. In the present example, this means keeping the lawful (anti-)correlation at the cost of sacrificing *B*'s being spin down.

[13] *Mutatis mutandis*, the following distinctions could still be made if this assumption were dropped. If $\tau_P$ were allowed to be a non-point-sized region of space-time then, in what follows, the phrases 'the absolute past of $\tau_P$' and 'the absolute future of $\tau_P$' should be replaced, respectively, by the phrases 'the union of all those points outside of $\tau_P$ that are in the absolute past of at least one point in $\tau_P$' and 'the union of all those points outside of $\tau_P$ that are in the absolute future of at least one point in $\tau_P$' (cp. Finkelstein [1999], p. 291).

[14] On these two options, compare (Bigaj [2004], [2006], pp. 185–90) and (Finkelstein [1999]). Other frame-invariant readings of Analysis 1 are possible. For instance, one might read it so that it requires us to hold fixed what lies to the past of a given Cauchy surface cross-cutting $\tau_P$'s absolute past close to $\tau_P$ (cp. Maudlin [2007], pp. 22–23; cp. also Stapp [2001]). The existence of further frame-invariant readings only strengthens our later diagnosis that Analysis 1 is ambiguous under relativity. For further discussion of temporal notions in relativity, see (Stein [1991]).

[15] It isn't the only reason. Another, related reason, is the relativistic prohibition upon a causal process connecting the measurement outcomes. Since standard CACs don't make the existence of a connecting process necessary for causation, the absence of such a process doesn't in itself prevent CACs from implying that the measurement outcomes are causally related. On the other hand, theories of causation that *do* take a connecting causal process to be necessary (e.g., Salmon [1984], [1994], [1997]; Dowe [2000]) straightforwardly imply that the case is *not* one of causation. Consequently, they can't do justice to our apparent ambivalence about such cases.

[16] As was the case with the frame-invariant readings of Analysis 1 that we discussed in Subsection 3.3, the following remarks can be generalized to space-time regions that aren't point-sized; cp. footnote 13.

[17] It was noted in footnote 14 that Finkelstein ([1999]) and Bigaj ([2004], [2006], Ch. 5) discuss the comparative merits of a pair of truth-conditions for counterfactuals that are closely related to the frame-invariant readings of Analysis 1. Those truth-conditions are also closely related to the $w_{\mathrm{AP}}$ and $w_{\mathrm{AFC}}$ approaches. Finkelstein and Bigaj don't seek to derive these truth-conditions from Lewis's priorities (1)–(4), however. Indeed, the context of their discussion differs from that of ours since, unlike Lewis, they disallow nomologically impossible worlds and hence disallow worlds involving violations of law (see Finkelstein [1999], pp. 289, 293–4; Bigaj [2004], pp. 7–8, [2006], pp. 97–101). The same restriction is in play in the debate about Stapp's ([1997]) proof; see especially (Shimony and Stein [2003], p. 501).

[18] Note that there is no need for an additional miracle to ensure that *A* isn't measured spin up in the first place. Since *A* had a chance of ½ of being measured spin down, the non-occurrence of its being spin up can be brought about without a miracle by simply having it be measured spin down.

[19] For ease of exposition, we assume that phrases of the form 'the closest worlds where such-and-such is the case' are well-defined, which requires the so-called Limit Assumption to hold (see Lewis [1973b], pp. 19–21). Nothing hinges on this, however.

[20] The state $(|\uparrow\uparrow\ldots\uparrow> + |\downarrow\downarrow\ldots\downarrow>)/\sqrt{2}$ for *n* particles yields an example of such a case. Suppose that an appropriate measurement is conducted on one of the particles, which in fact is spin up. Then, at a world at which this particle is instead measured spin down, it appears that $n - 1$ miracles are needed to hold fixed the remaining $n - 1$ particles in their original spin up state (cp. Lewis [1986a], p. 56).

[21] Bi-directional causal relations will be discussed further in the next section.

[22] Nor does our indecision seem to have anything to do with a possible epistemic uncertainty about how many particles are entangled, and consequently how many small miracles would be required.

[23] The question of whether Analysis 1 can be derived from a fully general semantics that is adequate to the ordinary language counterfactual (as Lewis claims Analysis 2 is) isn't one that need concern us. Even if Analysis 1 bears only a loose connection to the semantics appropriate to the ordinary language counterfactual, this doesn't tell against CACs that appeal to it. As Collins et al. point out concerning this issue, '[t]he counterfactual analyst should not worry: After all, she is doing the metaphysics of causation, and not the semantics of some fragment of English' ([2004], p. 9).

[24] Russell ([1913]) famously argued for an affirmative answer to this question. For recent discussion, see the papers collected in (Price and Corry [2007]).

[25] One could alternatively try to build a revisionary theory of causation on Analysis 2. This approach doesn't seem very promising, however. As we saw in Section 4, Analysis 2 yields definite causal verdicts when combined with standard CACs, but those verdicts depend upon how many particles are entangled, and it seems doubtful that a theory's causal verdicts ought to depend upon this.

[26] It is possible to forestall a possible objection to bi-directional causation voiced by Kistler ([2006], pp. 48–9). Suppose that events $e_1$ and $e_2$ cause each other, and that causation is transitive: if $e_1$ causes $e_2$ and $e_2$ causes $e_3$, then $e_1$ causes $e_3$. Then $e_1$ causes *itself* (assume $e_1 = e_3$), which may seem an unacceptable result. In response, we suggest that at most the following restricted transitivity principle, which is similar in spirit to the standard requirement

that cause and effect be distinct events, is true: if $e_1$ causes $e_2$, $e_2$ causes $e_3$, and $e_1$, $e_2$, and $e_3$ are pairwise distinct events, then $e_1$ causes $e_3$. This principle, together with the claim that there is bi-directional causation between our measurement outcomes, doesn't imply that each measurement outcome is a cause of itself.

## References


Bigaj, T. [2004]: 'Counterfactuals and Spatiotemporal Events', *Synthese*, **142**, pp. 1–19.

Bigaj, T. [2006]: *Non-Locality and Possible Worlds*, Heusenstamm: Ontos Verlag.

Butterfield, J. [1989]: 'A Space-Time Approach to the Bell Inequality', in J. T. Cushing and E. McMullin (eds), 1989, *Philosophical Consequences of Quantum Theory: Reflections on Bell's Theorem*, Notre Dame, IN: University of Notre Dame Press, pp. 114–44.

Butterfield, J. [1992a]: 'David Lewis Meets John Bell', *Philosophy of Science*, **59**, pp. 26–43.

Butterfield, J. [1992b]: 'Bell's Theorem: What it Takes', *British Journal for the Philosophy of Science*, **43**, pp. 41–83.

Butterfield, J. [2007]: 'Stochastic Einstein Locality Revisited', *British Journal for the Philosophy of Science*, **58**, pp. 805–67.

Collins, J., Hall, N. and Paul, L. A. (eds) [2004]: *Causation and Counterfactuals*, Cambridge, MA: MIT Press.

Dowe, P. [2000]: *Physical Causation*, Cambridge: Cambridge University Press.

Finkelstein, J. [1999]: 'Space-Time Counterfactuals', *Synthese*, **119**, pp. 287–98.

Hall, N. [2004]: 'Two Concepts of Causation', in J. Collins, N. Hall and L. A. Paul (eds), 2004, *Causation and Counterfactuals*, Cambridge, MA: MIT Press, pp. 225–76.

Hitchcock, C. [2004]: 'Do All and Only Causes Raise the Probabilities of Effects?', in J. Collins, N. Hall and L. A. Paul (eds), 2004, *Causation and Counterfactuals*, Cambridge, MA: MIT Press, pp. 403–17.

Hofer-Szabó, G. [2011]: 'Bell(δ) Inequalities Derived from Separate Common Causal Explanation of Almost Perfect EPR Anticorrelations', *Foundations of Physics*, **41**, pp. 1398–413.

Kim, J. [1976]: 'Events as Property Exemplifications', in M. Brand and D. N. Walton (eds), 1976, *Action Theory*, Dordrecht: Reidel, pp. 159–77.

Kistler, M. [2006]: *Causation and Laws of Nature*, London: Routledge.

Lewis, D. [1973a]: 'Causation', *Journal of Philosophy*, **70**, pp. 556–67.

Lewis, D. [1973b]: *Counterfactuals*, Oxford: Blackwell.

Lewis, D. [1979]: 'Counterfactual Dependence and Time's Arrow', *Noûs*, **13**, pp. 455–76.

Lewis, D. [1986a]: 'Postscripts to "Counterfactual Dependence and Time's Arrow"', in his *Philosophical Papers*, Vol. 2, 1986, New York: Oxford University Press, pp. 51–66.

Lewis, D. [1986b]: 'Postscripts to "Causation"', in his *Philosophical Papers*, Vol. 2, 1986, New York: Oxford University Press, pp. 171–213.

Lewis, D. [1986c]: 'Events', in his *Philosophical Papers*, Vol. 2, 1986, New York: Oxford University Press, pp. 241–69.

Lewis, D. [2004]: 'Causation as Influence', in J. Collins, N. Hall and L. A. Paul (eds), 2004, *Causation and Counterfactuals*, Cambridge, MA: MIT Press, pp. 75–106.

Maudlin, T. [2002]: *Quantum Non-Locality and Relativity*, 2nd ed., Oxford: Blackwell.

Maudlin, T. [2007]: *The Metaphysics Within Physics*, Oxford: Oxford University Press.

Price, H. and Corry, R. (eds) [2007]: *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*, Oxford: Clarendon Press.

Russell, B. [1913]: 'On the Notion of Cause', *Proceedings of the Aristotelian Society*, **13**, pp. 1–26.

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

Salmon, W. [1994]: 'Causality without Counterfactuals', *Philosophy of Science*, **61**, pp. 297–312.

Salmon, W. [1997]: 'Causality and Explanation: A Reply to Two Critiques', *Philosophy of Science*, **64**, pp. 461–77.

Shimony, A., and Stein, H. [2003]: 'On Quantum Non-Locality, Special Relativity, and Counterfactual Reasoning', in A. Ashtekar, R. S. Cohen, L. Divarci, D. Howard, J.

Renn, S. Sarkar, P. Schröter and A. Shimony (eds), 2003, *Revising the Foundations of Relativistic Physics: Festschrift in Honor of John Stachel*, Dordrecht: Kluwer, pp. 499–521.

Skyrms, B. [1984]: 'EPR: Lessons for Metaphysics', *Midwest Studies in Philosophy,* **9**, pp. 245–55.

Stapp, H. [1997]: 'Nonlocal Character of Quantum Theory', *American Journal of Physics*, **65**, pp. 300–3.

Stapp, H. [2001]: 'Response to "Comment on 'Nonlocal Character of Quantum Theory''", *American Journal of Physics*, **69**, pp. 854–59.

Stein, H. [1991]: 'On Relativity Theory and Openness of the Future', *Philosophy of Science*, **58**, pp. 147–67.

Woodward, J. [2003]: *Making Things Happen*, New York: Oxford University Press.