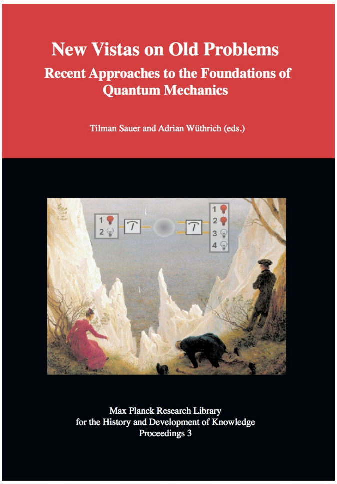
Max Planck Research Library for the History and Development of Knowledge
Proceedings 3

Matthias Egg:

Causal Realism in the Context of Bell-type Experiments



In: Tilman Sauer and Adrian Wüthrich (eds.): *New Vistas on Old Problems : Recent Approaches to the Foundations of Quantum Mechanics*
Online version at `http://edition-open-access.de/proceedings/3/`

ISBN 9783844242843
First published 2013 by Edition Open Access, Max Planck Institute for the History of Science under

Printed and distributed by:
Neopubli GmbH, Berlin
`http://www.epubli.de/shop/buch/26987`

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at `http://dnb.d-nb.de`

Causal Realism in the Context of Bell-type Experiments

Matthias Egg

Abstract. *After introducing the main idea of causal realism and discussing one of the key motivations for this position, I will review an argument by Tim Maudlin to the effect that there is superluminal causation in Bell/EPR-type experiments. I will then compare the concepts of causation used by causal realism and by Maudlin, in particular with respect to the importance they attach to practical controllability or manipulability of the causes. In conclusion, I will sketch how the causal realist can react to the impact of the EPR case.*

1 Introduction: Why Causal Realism?

A central question in the debate on scientific realism concerns the validity of inferences to the best explanation (IBE). Most strands of realism, whether they admit it or not, rely on this form of inference in one way or another. Accordingly, IBE has been a major target of antirealist criticism. The most basic line of criticism is that IBE commits the fallacy of *affirming the consequent*: if x explains y and y is true, it does not follow that x is true. Realists have, roughly speaking, responded to this charge in two ways. The first one, which I will not discuss here, consists in claiming that the theoretical virtues which mark out one explanation as *the best* can serve as an argument for the truth of this explanation (Psillos, 1999, ch. 8). The second realist response focuses on a specific class of scientific explanations, namely causal explanations, and claims that the particular character of the causal relation allows us to infer the *explanans* from the *explanandum*. If we explain some observed phenomenon y by saying that it was brought about by a cause x, it is legitimate to conclude that x really occurred. This is the kind of scientific realism that I call *causal realism*.

But how are we to understand this relation of *bringing about* between x and y? More precisely: how is causality supposed to do the job the realist wants it

T. Sauer, A. Wüthrich (eds.), *New Vistas on Old Problems. Recent Approaches to the Foundations of Quantum Mechanics*, Edition Open Access, 2013.

to do? As a first approximation, one might think that if x causes y, then x is a necessary condition for y, so that, whenever we observe y, we can be sure that x occurred as well. The inadequacy of this proposal becomes obvious once we consider how causal reasoning works in actual science. If, for example, particle physicists want to test whether some process x really occurs, they typically try to detect the products of x. More precisely, they calculate what kind of signal the process x should produce in their detector, and then they look for this signal (let us denote it by y) in the experimental data. Now typically, the mere occurrence of y by no means establishes the reality of x, because there are usually some alternative ways in which a signal of type y could have been produced. This is what physicists call "background." A case for x will only be made if there is a part in the counting rate for y that cannot be attributed to background. It will then in general be false to say that whenever y occurs, x has occurred as well. But it will be true that at least in some cases (though we may not know in which ones), y would not have occurred if x had not occurred. The truth of this counterfactual statement for at least some tokens of the event type y seems to be an essential part of what it means for x to cause y (regardless of the difficulties that a general account of causation in terms of counterfactuals may face). And the truth of this counterfactual establishes realism with respect to x.

Having thus tied causal realism to the truth of certain counterfactual statements, I might seem to have blurred the distinction between causal and theoretical realism. For it is not unique to causal explanations that they support certain counterfactuals. The same is true for theoretical explanations; they contain laws, and an essential aspect of lawhood is the property of supporting counterfactuals. But these counterfactuals are not of the right sort to lend the same kind of support to theoretical realism that causal statements lend to causal realism. Here is why: laws support counterfactual claims concerning their instances. For example, Boyle's law supports claims like "if I reduced the volume of this gas at constant temperature, its pressure would increase." But what theoretical realism aspires to establish by means of IBE is not the truth of a singular statement but of the law itself. In order to achieve this, a claim of the following form would be needed: "If law L did not hold, phenomenon y would not occur." This is actually not just a counter*factual*, but a counter*legal* statement, and the mere fact that L explains y implies nothing about the truth of such a statement. What it implies is that L is part of a sufficient condition for the occurrence of y, in the sense that L, conjoined with some initial conditions, allows us to derive a statement describing y. But using L in a theoretical explanation involves no speculation about what would happen if L did not hold. By contrast, as we have seen above, citing an entity x in a causal explanation of y essentially involves a claim about what would (or would not) have happened had x not occurred. In other words: the counterfactual statements that give rise to causal realism are an integral part of causal explanations. As a consequence, causal explanations are more closely linked to realism than theoretical explanations, and this is one motivation for causal realism.

To end this introduction, I briefly mention two other motivations for causal realism. One is that causal realism seems to be a promising strategy against a recent objection to scientific realism, introduced by Kyle Stanford and known as the *problem of unconceived alternatives*.[1] The other is that causal realism can be profitably combined with ontic structural realism, spelling out the latter as a metaphysics of causal structures (Esfeld, 2009).

2 The Argument for Superluminal Causation

When we ask whether there is superluminal causation in Bell/EPR-type experiments, a discussion of what exactly we mean by causation seems inevitable. I will touch on one aspect of this discussion in the next section, but first I will review an argument in favor of superluminal causation, which claims to involve only the most uncontroversial application of the notion of causation. Furthermore, it claims to hold independently of which particular solution to the quantum measurement problem one happens to prefer. The argument is from chapter 5 of Tim Maudlin's book *Quantum Non-Locality and Relativity*, an updated third edition of which has just recently appeared (Maudlin, 2011).

Maudlin starts by specifying a sufficient condition for a causal implication between two events: "The local physical events A and B are causally implicated with one another if B would not have occurred had A not (or vice versa)" (Maudlin, 2011, 117). This condition bears some resemblance to the counterfactual claim discussed in Section 1, but there is an important difference in the focus of inquiry: in Section 1, we started from an observable event y and asked about the reality of its unobservable cause x. Here, A and B are both observable events (typically the outcomes of measurements) and the question is whether there is a causal link between them.

[1] Chakravartty (2008, sec. 4) in response to Stanford (2006).

Obviously, the fact that A and B are causally implicated with one another in this sense does not yet imply that either A caused B or vice versa. If two television sets are tuned to the same program, it is correct to say that a certain picture would not have appeared on the first screen, had it not appeared on the second one. But it is not the case that the appearance of the picture on one of the screens caused the appearance on the other. Instead, there is a common cause for the two events, namely the signal sent out by the broadcasting company.

In the context of EPR-experiments, it is very natural to think that the observed correlations are due to a common cause, since these experiments typically involve particles coming from a single source, detected at different locations. Since the particles do not travel faster than the speed of light, the event of their emission at the source lies in the backward light cones of both the detection events. Therefore, if the emission could serve as a common cause explanation for the correlations, there would be no need for superluminal causation. So in order to argue for superluminal causation, we do not only need to show that two space-like separated events A and B are causally implicated with one another, but also that the implication cannot be traced to an event situated in the backward light cones of A and B. This is captured by Maudlin's *sufficient condition for superluminal influences:*

(SI) [G]iven a pair of space-like separated events A and B, if A would not have occurred had B not occurred even though everything in A's past light cone was the same then there must be superluminal influences. (Maudlin, 2011, 118)

It is obvious from the context that by "influences" Maudlin hear means "causal influences." Notice that the claim is not that there is a direct causal influence from either A to B or B to A. The causal connection between A and B may be due to a common cause C, but the condition (SI) states that C must lie outside A's backward light cone. But this is to say that there is superluminal causation between C and A. So whether we opt for direct causation between A and B or for some common cause, in either case there is superluminal causation.

How do we evaluate a counterfactual statement like the one in (SI), in order to decide whether (SI) is actually fulfilled in the context of EPR-experiments? Maudlin's answer is that "if we have gotten the laws of nature right, then we can know about at least some unrealized possibilities. Given a set of laws we may be able to evaluate counterfactuals, and thereby to discern some causal connections" (Maudlin, 2011, 120). At this point, a contradiction with causal realism might seem to arise, since, as argued in the introduction, causal realism maintains that our knowledge of the laws of nature is significantly less secure than our knowledge of causes. But if the evaluation of a causal claim like (SI) depends on a knowledge of certain laws of nature, then it seems that, contrary to what the causal realist believes, laws are epistemically prior to causes. However, this seeming contradiction can be resolved by distinguishing *fundamental* from *phenomenological* laws. It is only the former that arouse the causal realist's suspicions, because their acceptance depends crucially on their explanatory virtues, and, as discussed above, L's explaining y does not imply L's truth. By contrast, phenomenological laws derive their support from the simple fact that they accurately describe what is observed in experiments. The causal realist can endorse laws of this type wholeheartedly, and nothing more is required here. Consider, for example, the first part of (SI), the claim that "A would not have occurred had B not occurred." No deep theory is needed to justify this claim. Once we accept that there is a systematic correlation between the measurement outcomes in the left and the right wings of an EPR experiment ("systematic" in the sense that it can be expressed by a phenomenological law), we may infer that in at least some cases, the left outcome would have been different, had the right outcome been different.[2]

A somewhat more detailed treatment is needed to assess the second part of (SI), namely the claim that even if we held fixed everything in the past light cone of A (or B), the correlation between A and B would remain intact. But even here, the evaluation of the counterfactual does not depend on any specific theory. It only has to take into account that the measurement process which gives rise to the events A and B can be either deterministic or (irreducibly) stochastic. The two cases require two different treatments, but the result will be the same.

In the *deterministic* case, the assumption that a common cause located in the intersection of the backward light cones of A and B is responsible for the correlation implies a Bell-type inequality. The experimentally well-confirmed violations of such inequalities rule out any local-deterministic common cause model

[2] Maudlin formulates his argument in terms of perfect correlations, and in this case it is always true that the left outcome would have been different had the right outcome been different. This is of course highly idealized. Although I do think that experiments can in principle provide warrant for even idealized phenomenological laws, I will not argue for this here. It seems to me that Maudlin's argument goes through even with imperfect correlations, as long as they are assured to be non-accidental (and only the most radical sceptic will doubt that this latter fact can be established experimentally). Nevertheless, I will below assume perfect correlations whenever this simplifies the argument.

for the EPR correlations. In other words, assuming a deterministic measurement process, the correlation between the events A and B cannot be attributed to any causal factor located in the intersection of their past light cones. Thus (SI) is essentially[3] satisfied in this case. That the same is true for *indeterministic* models can most easily be seen in the case of a perfect correlation between A and B. If the measurement process that leads to A is truly stochastic, it could have come out differently even if its complete backward light cone remained unchanged. But had A come out differently, so would B (and vice versa), as required by (SI).

Since (SI) holds for deterministic as well as indeterministic models, it follows that the existence of superluminal causation is established independently of any specific approach to the measurement problem. Maudlin concludes: "Reliable violations of Bell's inequality need not allow superluminal signaling but they do require superluminal causation" (Maudlin, 2011, 141). This contrast brings us to the topic of the next section, namely signaling and how it relates to the concept of causation.

3 Causation, Manipulability, and Signaling

One reaction to the verdict of the previous section is to ask whether Maudlin has rigged the game by helping himself to too weak a notion of causation. If there is a causal relation between A and B, should it not at least in principle be possible to bring about a variation in B by manipulating A? And if so, should it not be possible to use that manipulation to send a signal from A to B, thereby violating some no-signaling theorem? Maudlin discusses this question in a section entitled "But is it causation?" (Maudlin, 2011, 135–141) There he argues that the exploitability or controllability of the causal relation should not be part of the concept of causation and that no-signaling should therefore not be taken to imply no-causing: "In general if one adds control of one variable to a counterfactual-supporting connection one gets signaling, but the addition is strictly irrelevant to the existence of the causal connection" (Maudlin, 2011, 137).

No matter whether or not one agrees with this characterization, one might ask at this point if the issue is relevant at all. It certainly is interesting to learn about these non-local dependencies, but does it make any difference whether we call

[3] There is a small argumentative gap here, because (SI) requires holding fixed the *entire* past light cone of A and not just the part that overlaps with the past light cone of B. See (Maudlin, 2011, 122) for an argument closing this gap.

them causal or not? Well, from the perspective of causal realism, it does make a difference which structures are causal and which are not, because this affects the decision on how far the realist commitment should extend. Furthermore, practical manipulability has played an important role in some of the arguments for causal realism. Consider, for example, Ian Hacking's famous pronouncement about electrons: "So far as I'm concerned, if you can spray them then they are real." (Hacking, 1983, 23) It thus seems worthwhile to consider in more detail how the lack of manipulability affects realism in the context of EPR experiments.

As we saw in the last section, the argument for superluminal causation (in Maudlin's sense) is independent of any specific choice of theoretical model and it is backed by strong experimental evidence. So far, the story is perfectly acceptable to the causal realist. But what exactly does this commit him to? To the reality of superluminal influences, of course, but what kind of influences and between which relata? Maudlin's argument can only be as general as it is by refusing to answer these questions. For illustration, let us look at two possible ways to account for EPR correlations.[4]

The most obvious option is to postulate a direct superluminal influence from A to B (or vice versa). Apart from being faster than light, such influences would be unusual in yet other ways: their strength does not seem to diminish when the distance between A and B is increased[5] and the influence is discriminating in that it only affects particles that have previously interacted with each other.

The second option tries to avoid such action at a distance by denying that there are two entities, one in each wing of the experiment, influencing each other across a space-like interval. Rather there is one single quantum structure that brings about events A and B. But of course we still have a superluminal influence: event A, for example, is caused by the whole, non-separable quantum structure which spans both wings of the experiment, so A is influenced by something which is not confined to its past light cone.

I will not enter into a discussion about which of these models is preferable or less objectionable. The point is that they are both compatible with what experiments tell us about EPR-like arrangements. The situation is similar to one that appears frequently in empirical research based on statistics: we observe a corre-

[4]For more examples, see (Suárez, 2007).

[5]A referee has helpfully pointed out that this is actually not so unusual; the strong nuclear force, as described by quantum chromodynamics, even *increases* with increasing distance between the interacting particles. This is true, but such behavior is restricted to subatomic distances. By contrast, EPR correlations have been shown to extend over distances of several kilometers.

lation between two variables, but we do not know whether this correlation is due to a direct causal influence from one variable to the other or to a common cause of the two variables. This analogy allows us to see why causal realism places such emphasis on the practical manipulability of alleged causes. For this is precisely what enables us in many cases to discern the causal structure that holds between the variables. Consider once again the example of the two TV sets from the previous section. The fact that we cannot, by manipulating the image on the first screen, influence what appears on the second screen, strongly favors the common cause hypothesis over the hypothesis that one of the images causes the other.

The controllability of a causal factor, therefore, has an important epistemic significance over and above the obvious fact that a controllable causal factor may open the way for technical applications. In the absence of controllability, we might have strong empirical evidence for the existence of some causal relations (as the argument in the previous section shows), but we do not have this type of evidence for claims about the precise causal structure of the situation.

4 Conclusion: Towards a Bell-informed Causal Realism

Although the result of the last section justifies causal realism's insistence on controllability/manipulability, it seems to have a rather devastating implication for this type of realism in the context of Bell-type experiments. If causal realism can only get off the ground when manipulation of the relevant factors is possible and if the no-signaling theorem prohibits such manipulation in the EPR cases, then it seems that causal realism is simply irrelevant in this context. I will now sketch two possible strategies to avoid this conclusion.

The first strategy starts from the observation that the no-signaling theorems are theoretical results, so they are only as well confirmed as the theories from which they are derived. Of course, standard quantum mechanics is extremely well confirmed, but by itself, it does not provide us with a satisfactory account of what happens in a measurement. Since this is precisely one of the ingredients of the EPR puzzle, it is at least possible that a theory which solves the measurement problem will yield predictions on signaling that differ from those of standard quantum mechanics. One might object that this is not a very plausible scenario, given that there are presently no empirical indications for a violation of no-signaling. But the fact that something has not been possible up to now does not imply that it should be impossible in principle. The history of science contains many examples of clever experimenters who gained access to domains that were previously thought to be inaccessible in principle.

But even if future inquiry should lead to progress along these lines, some ambiguities in the causal structure of EPR correlations will probably remain unresolved. In the previous section I merely discussed a very coarse-grained distinction between a common cause and a direct cause model, which, if we presuppose controllability of the measurement events, would be comparatively easy to resolve experimentally. However, each of these generic options is compatible with a number of more fine-grained models concerning the precise character of the causal influence, and experiments may fail to distinguish between these models. This point reflects a general dilemma for causal realism: on the one hand, the causal realist wishes to limit his commitments to claims for which there is direct experimental evidence, on the other hand, such evidence may not be available for detailed claims about the causes with which causal realism is concerned.

The second strategy for an improved causal realism seeks to deal with this dilemma in a constructive manner, not really seeing it as a dilemma at all, but as a hint on how to differentiate between different grades of commitment within one's realism. On this view, the two horns of the dilemma correspond to *two kinds of warrant* (called *causal* and *theoretical* by Suárez (2008) and Egg (forthcoming)), which can be ascribed to scientific claims. Claims that are causally warranted form the hard core of the realist's commitment, because they are as secure as any empirical claim can be. In particular, they can be defended against any kind of antirealist criticism (though not against radical skepticism, of course). The price to pay for this security is a lack of specificity. We saw an example of this in the discussion of superluminal influences: there is no reasonable doubt that there are such influences, in other words, we have causal warrant for their existence, but this commitment does not include any details about their precise nature. However, causal realism does not advocate complete agnosticism with regard to these details. In the absence of direct experimental evidence, causal realism can draw on the resources of standard scientific realism and evaluate the theoretical warrant for the different models that are compatible with experience, in order to arrive at a more detailed picture than what the evaluation of causal warrant by itself would yield.

The work of formulating such a detailed picture for the EPR experiments remains to be done. But whatever the precise outcome will be, the approach of causal realism will have the advantage of transparently delimiting the parts of the picture to which we should be strongly committed as opposed to the more speculative parts. The fascinating thing about EPR is that even the causally warranted core part of the picture contains the potential for a serious clash with common sense and special relativity.

References

Chakravartty, A. (2008). What you don't know can't hurt you: realism and the unconceived. *Philosophical Studies*, 137:149–158.

Egg, M. (forthcoming). Causal warrant for realism about particle physics. *Journal for General Philosophy of Science*.

Esfeld, M. (2009). The modal nature of structures in ontic structural realism. *International Studies in the Philosophy of Science*, 23(2):179–194.

Hacking, I. (1983). *Representing and Intervening*. Cambridge University Press, Cambridge.

Maudlin, T. (2011). *Quantum Non-Locality and Relativity*. Wiley-Blackwell, Chichester, 3rd edition.

Psillos, S. (1999). *Scientific Realism: How Science tracks Truth*. Routledge, London.

Stanford, P. K. (2006). *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. Oxford University Press, New York.

Suárez, M. (2007). Causal inference in quantum mechanics: a reassessment. In Russo, F. and Williamson, J., editors, *Causality and Probability in the Sciences*, pages 65–106. London College, London.

Suárez, M. (2008). Experimental realism reconsidered: How inference to the most likely cause might be sound. In Hartmann, S., Hoefer, C., and Bovens, L., editors, *Nancy Cartwright's Philosophy of Science*, pages 137–163. Routledge, New York.