TOMASZ BIGAJ

The Common Cause Explanation for Quantum Correlations of the EPR Type

Abstract. In the article van Fraassen’s argument against the possibility of the common cause explanation of EPR correlations is analyzed. It is pointed out that the crucial assumption of this argument is given in an ambiguous way, and that only one of its interpretations leads to the above conclusion. Several other omissions and misleading statements in van Fraassen’s analysis are brought into considerations. At the end the corrected version of the original argument is given and its negative result is finally accepted.

Key words: Bell’s theorem, common cause, EPR correlation, probability

1. The Philosophical Importance of the Bell Theorem

It is unquestionable truth that Bell’s theorem has profound implications concerning fundamental characteristic of the physical world. However, what exactly these implications are remains to be the subject of controversy among philosophers of science. The most common interpretation of Bell’s result is that it excludes, or at least makes extremely unlikely, the thesis of realism—in other words the hidden variables hypothesis (cf. for example Hughes 1989, p. 170 ff). On the other hand, some theorists claim that Bell’s theorem undermines the thesis of locality only, and that it has nothing to do with the hidden variables issue (Maudlin 1994, pp. 19–20). More modest approaches have it that the Bell inequality leaves us with the choice: either abandon realism, or accept nonlocality in its strongest form, implying that a change in property of one system can instantaneously change some property of a distant system (Redhead 1987). On top of this, another general epistemological

1 Another recent analysis of implications of the Bell theorem deserves to be mentioned here. J. Hawthorne and M. Silberstein claim that they were able to derive the strongest possible consequences from the Bell theorem, with the help of the weakest assumptions, for example without assuming perfect correlations between results of measurements on distant systems (Hawthorne, Silberstein 1995). They claim that the Bell inequality implies both dependence of measured properties on the
consequence appeared to be derivable from the Bell inequality. In 1982 Bas van Fraassen came up with an ingenious argument, showing that any attempt to causally explain EPR-type correlations, which respects the limitations implied by the STR, is doomed to be in conflict with experience. More specifically, van Fraassen proved that—under certain conditions—the hypothesis stating that there must be a common cause for every correlation between results of measurements on distant particles in an entangled quantum state leads inevitably to the Bell-Wigner inequality, which is known to be empirically falsified. The paper containing this result was first published in *Synthese* (van Fraassen 1982), then it was reprinted in 1989 with postscript in (Cushing, McMullin 1989), and finally included as a main part of Chapter 4 *The Empirical Basis of Quantum Theory* in van Fraassen’s book *Quantum Mechanics: An Empiricist View* (van Fraassen 1991a). From this triple occurrence it can be seen that the author himself must have attached great importance to this particular piece of his work. However, to my best knowledge the argument presented there, although widely acknowledged (for example it was referred to in the above-mentioned book (Hughes 1989, p. 246), and mentioned in Hawthorne, Siberstein 1995), has not been subjected to satisfactory scrutiny. In this paper I would like to focus in details on van Fraassen’s derivation of the Bell inequality from the hypothesis of the common cause. It appears that his argument contains serious omissions and surprising slips. For example I will show that the key assumption of the argument is given in an ambiguous way, and that the validity of the entire reasoning depends on the decision which of its interpretations is proper. My main goal will be to fill some of the lacunas in van Fraassen’s reasoning and see if there is any chance to block its conclusion, which is devastating for everyone who believes in causal order of the physical world. In my subsequent analysis I will be referring to the last version of van Fraassen’s paper, printed in his book (1991).

The starting point of van Fraassen’s argument is the general definition of the common cause $C$ for two correlated events $A$ and $B$, supposedly based on Reichenbach’s proposal which can be found in (Reichenbach 1956). The definition, as stated by van Fraassen, consists of three requirements (p. 83):

1. $C$ precedes $A$ and $B$,

type of measurement, and nonlocality or holism. The first part seems to be close to the negation of realism, or the thesis that there are objective properties of systems (or at least objective dispositions) revealed upon measurement. So it appears as if finally realism were undermined independently of the issue of locality. However, it is not the case. A close inspection reveals that the authors broadened the notion of measurement dependence so that it covers cases in which the result obtained on one system depends on the type of a measurement performed on the other, distant system. But this is exactly what is usually meant by a non-local influence. If the authors had succeeded in showing that Bell’s theorem implies that the results of any measurement must depend nontrivially on the type of this particular measurement, then this result would have been a real breakthrough. But in fact they still face the usual dilemma: in order to account for the empirical falsification of the Bell inequality, we can either claim that the results of measurements are nontrivially brought about by the type of a measurement performed, or we can shield the thesis of the independent reality of measurable property by invoking non-local disturbances caused by a distant measurement. Needless to say, the first option is still vulnerable to some kind of non-local influences, as the authors show in the further sections of their paper.
(2) \( P(A|C) > P(A|\text{not-}C) \) and \( P(B|C) > P(B|\text{not-}C) \),

(3) \( P(A \& B|C) = P(A|C)P(B|C) \).

For the time being we are leaving this definition uncontested, but we will have to return to the question of its tenability (and the faithfulness to Reichenbach’s own proposal) soon. For now we can only mention that according to van Fraassen the condition (1) is supposed to exclude some artificial causal “explanations” of the correlation between \( A \) and \( B \), for example by stipulating \( C = A \& B \). It is easily seen that such \( C \) fulfills the conditions (2) and (3), yet it cannot be treated as a proper common cause for \( A \) and \( B \).

2. Probability Functions and Probability Spaces

The general framework in which van Fraassen presents his EPR-like situation is quite clear and well known. We have two particles \( L \) and \( R \), three settings of experiments on each of them, and two possible outcomes of each experiment. The paradigmatic example of such a situation would be a measurement of polarization in three distinct directions for a pair of photons, with two possible outcomes “passed” or “absorbed”, or a measurement of three distinct components of spin of electrons, with two results “up” or “down” each. The proposition \( \text{ Lia } (Rjb) \) describes the fact that the experiment conducted on the particle \( L \) (\( R \)) in the setting \( i \) (\( j \)) gave the result \( a \) (\( b \)). The probability space considered by van Fraassen is made up of propositions of the form \( \text{ Lia } \& Rjb \), where \( i, j = 1, 2, 3 \) and \( a, b = 0, 1 \). Obviously, this space includes 36 distinct propositions. Now, in principle we could consider an unconditional probability function defined on all 36 propositions. In that case for example the probability \( P(L10 \& R20) \) would have the following empirical meaning: it would equal (as a limit) the relative frequency of the occurrence of the result 0 in the setting 1 for the particle \( L \) together with the result 0 in the setting 2 for the particle \( R \) in the long run of experiments. But note that this frequency involves in fact two separate phenomena: it depends on the (in principle) objective “disposition” to give certain results, as well as on our choice of settings. So, we could for example make any of those frequencies equal zero, simply by not choosing the particular setting for a measurement. One way to avoid this would be using a random mechanism to choose settings for each measurement, but a better way is to consider conditional probabilities and thus to separate objective facts from our decisions.

For this reason van Fraassen considers only probabilities which he calls “surface probabilities”. These are conditional probabilities of the form \( P(\text{ Lia } \& Rjb|Li \& Rj) \), where \( Li = Lj0 \lor Lj1 \) and \( Rj = Rj0 \lor Rj1 \) (proposition \( \text{ Lia } \) is in turn equivalent to the disjunction of the propositions \( \text{ Lia } \& Rjb \) for all indices \( j \) and \( b \)). In other words, we are interested only in probabilities of a given result of an experiment, conditionally on choosing its setting. It is quite obvious that surface probabilities can be extended to a total probability measure on the entire probability space, given the probability of the settings \( Li \& Rj \) (which van Fraassen calls a “choice weighting”, see pp. 87-88). We can prove this by appeal to the following theorem of the
probability calculus (assuming that the propositions $Lk \& Rl$ are mutually exclusive and jointly exhaustive):

$$P(Lia \& Rjb) = \sum_{Lk} P(Lia \& Rjb|Lk \& Rl) P(Lk \& Rl).$$

From this it follows that in order to calculate the unconditional probabilities on the set of propositions $Lia \& Rjb$, even when the probability distribution on $Li \& Rj$ is known to us, we should in principle be able to determine the conditional probabilities of the form $P(Lia \& Rjb|Lk \& Rl)$, where $i \neq k$ or $j \neq l$. Fortunately, because the proposition $Lk (Rl)$ is equivalent to $Lk0 \lor Lk1 (Rl0 \lor Rl1)$, and because we assume that the propositions of the form $Lia$ and $Ljb$ are mutually exclusive for $i \neq j$ (i.e. the equation $P(Lia \& Ljb) = 0$ holds for $i \neq j$), we can easily verify that the above conditional probabilities are equal to zero, and that the above-mentioned formula reduces to the following:

$$P(Lia \& Rjb) = P(Lia \& Rjb|Li \& Rj) P(Li \& Rj).$$

If the probability distribution over setting choices were uniform, i.e. if the condition $P(Li \& Rj) = P(Lk \& Rl)$ held for all $i, j, k, l$, the relation between surface probabilities and full probabilities would be even simpler: $P(Lia \& Rjb) = \frac{1}{9} P(Lia \& Rjb|Li \& Rj)$. Note that the unconditional probability of the form $P(Lia \& Rjb)$ still refers to the results obtained in the actual experiment, and not, as it might be tempting to assume, to the results that could have been obtained had the setting been chosen different. In other words, the truth of the proposition $Lia (Rjb)$ logically implies the truth of $Li (Rj)$, because $Lia$ is equivalent to the conjunction of two sentences: that we have chosen the setting $i$ and that the outcome of the measurement was $a$.

In connection with the above remark note that we could have introduced a slightly different kind of basic propositions. Namely we could have chosen the propositions of the form $Lia^* (Rjb^*)$, whose intended meaning would be that the actual value of an appropriate property in the setting $i (j)$ for the particle $L (R)$ equals $a (b)$. This reading doesn’t make any appeal to measurements performed or settings chosen. It has an objective character, and goes beyond what is actually measured in experience. The other way of reading the propositions $Lia^*$ would be in terms of counterfactual (modal) conditionals: “If the setting $i$ for the particle $L$ had been chosen, then the result of the measurement would have been $a$”. We can believe that at every moment the particle has a determined (probabilistic) disposition to give certain outcomes of different measurements. So, to explain it a little bit, we can interpret the probability $P(Lia^*)$ as the probability of the fact that the result would be $a$, if the measurement were performed in setting $i$ (even when no actual measurement has been performed, or when another setting has been chosen), whereas the conditional probability $P(Lia^*|Li)$ can be interpreted as the probability of the fact that the result of a hypothetical measurement in the setting $i$ would be $a$. **given that we have actually performed this measurement in the setting $i$**. Of course under such interpretation it is no longer valid that the proposition $Li$ (stating that the setting chosen is $i$) is equivalent to the disjunction $Li0^* \lor Li1^*$. In spite of
this fact the conditional probability \( P(Lia^* \& Rjb^*|Li \& Rj) \) still would be empirically tested. Moreover, it can be claimed that this probability is exactly the same as \( P(Lia \& Rjb|Li \& Rj) \). This should be clear when we assume the principle of faithful measurements, stating roughly that measurements reveal objective propensities of particles to give appropriate experimental results. However, in this situation the transition from surface probabilities to unconditional probabilities is not possible, even given a choice weighting, because we cannot be sure that the probabilities \( P(Lia^* \& Rjb^*|Lk \& Rl) \) for \( i \neq k \) or \( j \neq l \) reduce to nil (in fact, it is impossible to measure these probabilities empirically). So the probability \( P(Lia^* \& Rjb^*) \) will have to remain empirically undecidable.\(^2\)

Summing up we may say that in principle three different probability functions can be used in quantum mechanical description. Two of them would be unconditional, and one conditional. The unconditional functions can be defined either on propositions referring to actual outcomes, or on propositions about possible outcomes. The probability function defined on the counterfactual propositions of the type \( Lia^* \) can be called “deep probability”. We will continue to refer to the probabilities conditional on settings chosen as “surface probabilities”, whereas unconditional probabilities on the propositions stating the results of actual experiments can be named “experimental probabilities”, as they contain some information about how the experimenter decided to choose its measuring setups. Both experimental and surface probabilities can be determined empirically in the form of relative frequencies, whereas deep probability lies as it seems outside the scope of simple experimental verifications.

In spite of the last remark, it is possible to imagine an argument for the empirical decidability of deep probabilities of the form \( P(Lia^* \& Rjb^*) \). In order to consider this, we must again appeal to the equation \( P(Lia^* \& Rjb^*) = \sum_{i,j} P(Lia^* \& Rjb^*|Lk \& Rl) P(Lk \& Rl) \). However, what is the meaning of the conditional probability \( P(Lia^* \& Rjb^*|Lk \& Rl) \)? It can be read as follows: it is the probability of getting the result \( a \) for the particle \( L \) and the result \( b \) for the particle \( R \) in the counterfactual situation with the settings \( i \) and \( j \) respectively, given that the actual settings were \( k \) and \( l \). But then, what difference do the settings actually chosen make, if we are interested only in what would have been, had the setting been such-and-such? The actual choice of settings should not have any impact on the probability defined on counterfactual conditionals. In that way we could argue that all the probabilities of the form \( P(Lia^* \& Rjb^*|Lk \& Rl) \) should be equal. Therefore the initial equation would come down to: \( P(Lia^* \& Rjb^*) = P(Lia^* \& Rjb^*|Li \& Rj) \)

\(^2\) Let us point out some other differences between propositions of the form \( Lia \) and \( Lia^* \). For example, the truth of the former does not imply the truth of the latter. This is so because even if in the reality the measurement were performed and the result were \( Lia \), it nevertheless doesn’t justify the claim that if that experiment were performed anew, the result would be the same (this may suggest that in proper semantic description of the propositions \( Lia^* \) we should resort to many-valued logic). It justifies only the thesis that the probability of \( Lia^* \) must have been greater than zero. And similarly, from the truth of \( Lia^* \) obviously doesn’t follow that \( Lia \) is true, for we could have chosen a different setting. But if it holds that \( P(Lia|Li) = 1 \), then \( Lia^* \) should be true. Therefore we can claim the following equivalence \( P(Lia^*) = 1 \iff P(Lia|Li) = 1 \), which of course doesn’t imply that \( P(Lia|Li) = P(Lia^*) \).
the sum reduces to 1, which finally gives \( P(L_{ia}^* \& R_{jb}^*) = P(L_{ia}^* \& R_{jb}^*|L_i \& R_j) \). But this last equation, on the basis of our previous analysis, equates deep probabilities with surface probabilities! I suspect that when for example physicists talk in quantum mechanics, as they often do, about unconditional probabilities of getting certain results in certain measurements, they have a similar argument in minds. For now I will leave this issue at this stage, noting only that van Fraassen would probably oppose the move of equating all the conditional probabilities of the form \( P(L_{ia}^* \& R_{jb}^*|L_i \& R_j) \).³

3. Assumptions

The first two assumptions of van Fraassen’s argument are straightforward. The correlation between two results of the experiment carried out with the same settings should be perfect (as in the EPR situation with spin \( \frac{1}{2} \)) which means that if the result on the particle \( L \) is \( a \), the result of the same measurement on \( R \) must be \( 1-a \). This condition finds the following formulation:

\[ \text{I. } P(L_{ia} \& R_{ia}|L_i \& R_i) = 0 \]

Of course the equation (I) is valid for all values of \( i \) and \( a \). In the next step, van Fraassen formulates an assumption which, although not used directly in the derivation of the Bell inequality, plays an important role as a kind of heuristic justification for another premise in the argument. This assumption is dubbed “surface locality” and says roughly that the result of an experiment on one particle does not depend on the setting chosen for the distant partner. This fact is of course empirically testable, and it has actually been verified. We can write it down as follows:

\[ \text{II. } P(L_{ia}|L_i \& R_j) = P(L_{ia}|L_i) \]
\[ P(R_{jb}|L_i \& R_j) = P(R_{jb}|R_j) \]

In other words: only conditionalizing on the particle’s own setting matters. Now let us move to the central assumption of van Fraassen’s derivation, which introduces a hypothetical common cause, aiming at an explanation of the correlation (I). Van Fraassen postulates the existence of a hypothetical parameter \( A \) with values \( q \), associated with the particle source, and acting as a common cause in Reichenbach’s

³ Van Fraassen expressed on many occasions his deep suspicion towards using counterfactual conditionals in interpreting results of measurements in quantum mechanics (see for example van Fraassen 1991b, p. 122-125). His claim is that unrestricted usage of such counterfactuals may lead to a contradiction with the principles of quantum mechanics and with experience (via Bell’s theorem). However, the problems which van Fraassen points at are not necessarily connected with counterfactuals themselves, but with certain interpretations of them. For example, we must not accept the thesis of counterfactual definiteness, which says that one of the following sentences “If the measurement were done, the result would be \( a \)” is true. In fact, our proposition \( L_{ia}^* \) sometimes can be neither true, nor false – when the probability of obtaining the value \( a \) is between 0 and 1. So I would claim that we can safely resort to counterfactuals, but under the condition that we abandon the classical logical principle of bivalence.
sense—namely screening off statistically correlated events. The condition written by van Fraassen is apparently meant to be an instantiation of the requirement (3) from the general definition of the common cause.

III. \( P(Lia \& Rjb|Li \& Rj \& Aq) = P(Lia|Li \& Rj \& Aq) \; P(Rjb|Li \& Rj \& Aq) \)

But now an interpretative problem appears. In the original formulation there was only one common cause \( C \), yet here we have as many of them as there can be different values of the parameter \( A \) (because every fact that \( A \) takes some value \( q \) is a separate event). And besides, we have not only one correlation to explain but (in general) 36 of them. So how exactly should we understand the condition (III)? There seem to be at least two interpretations. In the weaker of the two, for all values \( i, j, a, b \) there is a value \( q \) such that (III) holds, which amounts to the fact that every correlation is causally explained by some event in the past. But according to the stronger interpretation, equation (III) should hold for all values of \( i, j, a, b \) and \( q \). It is impossible to decide between the two interpretations on the basis on van Fraassen’s text only, for he entirely omitted quantifiers in the formulation of his assumptions. Neither his remarks suggest which interpretation is proper. So we must decide on our own. First let us state explicitly both interpretations.

IIIa. For all \( i, j, a, b \) there is \( q \) such that
\[
P(Lia \& Rjb|Li \& Rj \& Aq) = P(Lia|Li \& Rj \& Aq) \; P(Rjb|Li \& Rj \& Aq)
\]

IIIb. For all \( i, j, a, b \) and \( q \)
\[
P(Lia \& Rjb|Li \& Rj \& Aq) = P(Lia|Li \& Rj \& Aq) \; P(Rjb|Li \& Rj \& Aq)
\]

Now, at first glance it seems that the condition IIIa is closer to the original definition which van Fraassen cited after Reichenbach. For it was only required there that in order to causally explain a correlation between two events, we must postulate one common cause \( C \). In the EPR case we have (as we noted) 36 correlations to explain, but it is not said that we must explain them “together”. The condition IIIb suggests that the fact that the parameter \( A \) has some value, explains all the correlation involved. But isn’t this assumption too strong? Wouldn’t it suffice to say that every correlation is causally explained at least by some value (and not necessarily the same) of the parameter \( A \)?

Let us leave these questions unanswered for the moment and quickly present the last two assumptions. The fourth assumption is a generalization of surface locality for the cases with fixed value of the parameter \( A \). Van Fraassen gives it in the following way.

IV. \( P(Lia|Li \& Rj \& Aq) = P(Lia|Li \& Aq) \)
\[
P(Rjb|Li \& Rj \& Aq) = P(Rjb|Rj \& Aq)
\]

It simply results from adding \( Aq \) to the probabilities in the requirements (II) as an additional condition. Van Fraassen’s last premise can be seen as a formal equivalent of the informal condition (1) from the definition of the common cause. It states that the parameter \( A \) has value \( q \) is statistically independent of the choice of settings for both experiments on \( L \) and \( R \). Because the parameter \( A \) is meant to characterize the particle source, this condition implies that there cannot be “backward causation” leading from the later choice of settings to the earlier state of the source, but
it also excludes the possibility, only a little bit less awkward, that the exact value of
the parameter can influence the experimenter’s choice or the mechanism used to
randomly select settings. Van Fraassen calls this assumption “hidden autonomy”
and presents it in the following form:

\[ V. \ P(Aq|Li \ & \ Rj) = P(Aq). \]

4. Derivation

Van Fraassen’s derivation of the Bell inequality uses four of the above assumptions
(excluding II). Essentially it consists of two simply steps. In the first step the thesis
of determinism is deduced. Determinism is the claim that every value \( q \) of the pa-
rameter \( A \) uniquely determines all the results of possible experiments on particles \( L \)
and \( R \). The formal exposition of this theorem is as follows:

For all \( i, j, a, b, q \)

\[ \text{(DET)} \quad P(Lia|Li \ & \ Aq) = 0 \text{ or } 1 \]
\[ P(Rib|Rj \ & \ Aq) = 0 \text{ or } 1. \]

Although the derivation of (DET) from the premises (I), (III) and (IV) seems to
be straightforward, it is important for our purposes to state it carefully, because it
will enable us to choose the right form of the condition (III). For it is only the
stronger condition (IIIb) and not the weaker (IIIa) which allows to derive (DET).
Van Fraassen begins with the perfect correlation condition (I):

\[ 0 = P(Lia \ & \ Ria|Li \ & \ Ri) \]

If the conditional probability is 0, then this value cannot change when we add any
extra element to the condition. Therefore we can add \( Aq \).

\[ 0 = P(Lia \ & \ Ria|Li \ & \ Ri \ & \ Aq) \]

And now time has come for assumption (III). Van Fraassen simply replaces the
right-hand side of the equation with the product of two probabilities.

\[ 0 = P(Lia|Li \ & \ Ri \ & \ Aq) \ P(Rib|Li \ & \ Ri \ & \ Aq) \]
\[ P(Lia|Li \ & \ Ri \ & \ Aq) = 0 \text{ or } P(Rib|Li \ & \ Ri \ & \ Aq) = 0 \]

Of course whichever of the probabilities is zero, the other must be one. Therefore
we have arrived at the conclusion:

\[ P(Lia|Li \ & \ Ri \ & \ Aq) = 0 \text{ or } 1 \]
\[ P(Ria|Li \ & \ Ri \ & \ Aq) = 1 \text{ or } 0 \]

or, using (IV)
\[
P(Lia|Li & Aq) = 0 \text{ or } 1
\]
\[
P(Ria|Ri & Aq) = 1 \text{ or } 0.
\]

Until this moment all versions of (III) seem to work. But let’s see exactly what we have proven. If we used the assumption (IIIa), our result would be only that for every setting \(i\) there is some value \(q\) of the parameter \(A\) such that the fact \(Aq\) uniquely determines results of measurements \textit{in this setting}. But this value of the parameter doesn’t have to determine results in all other settings. In other settings the above equations hold, but possibly for \textit{different values of \(A\)}. Therefore (DET) is not proven, because it says that for all values of \(A\) and for all settings the results are determined. In order to prove this, we must resort to the stronger interpretation (IIIb).

So it seems that the problem of the interpretation of (III) can be finally solved and that we can only accuse van Fraassen of not being rigorous enough in formally presenting his assumption (III) He should have explicitly stated “for all \(i, j, a, b, q\)” before the formula (III), and this is probably what he had in mind. After all, it is a common practice to omit general quantifiers in formulas with variables. But this is not so simple. Firstly, we must ask for the justification of the stronger version (IIIb), because as we noted, on the surface it is not directly implied by the conditions (1)–(3). And secondly, after close inspection of van Fraassen’s text we get the impression that he himself hesitates which version of (III) should be adopted. Some of his comments and remarks can be interpreted as assuming the weaker understanding of the common cause condition. In order to illustrate this, let us consider in some details his remark on p. 90 concerning the fact that without conditions (IV) and (V) condition (III) alone could be satisfied by some artificially chosen “common causes”. Van Fraassen claims that if we defined \(Aq\) as the proposition stating the actual outcomes of experiments in actual settings, condition (III) would be trivially satisfied.

But is it really so, especially if we interpret (III) as (IIIb)? First of all notice that the use of the phrases “actual settings” and “actual results” can be a little bit misleading. In van Fraassen’s argument there are no other settings chosen than the actual, and we cannot obtain other results except the actual ones. When he uses the word “actual” it might however suggest that there is a difference between the \textit{actual setting} and setting \textit{simpliciter}. He even strengthens this suggestion by using the index \(t\) (standing for “true”, I suppose) in the stipulation: \(Aq = Lia, & Rjb,\) apparently to distinguish this proposition from the simple \(Lia & Rjb\). But wait a second. How should we, for example, read the probabilities of the form \(P(Lia & Rjb|Li \& Rj)\)? It goes as follows: this is the probability of obtaining jointly the result \(a\) in setting \(i\) on \(L\) and the result \(b\) in setting \(j\) on \(R\), given that the \textit{actual} settings were respectively \(i\) and \(j\). It is obvious that we cannot conditionalize on other settings than actual (we can do it \textit{counterfactually}, but this can’t be expressed with the help of the conditional probabilities, but rather of propositions of the form \(Lia^*\)). Therefore there is no need (and no justification) to use any indices to emphasize the actuality of the settings and the results. The mere fact that we conditionalize on an event presupposes its actuality. Hence van Fraassen’s claim
amounts in fact to this that when we put \( A_q = L_{kc} & R_{ld} \) for any indices \( k, l, c, d \), then the condition (III) will be satisfied. But now we see that the justification of this claim depends on the interpretation of (III). The weaker condition (IIIa) would be of course satisfied, because it says only that for every proposition \( L_{ia} & R_{jb} \) there is some value \( q \) of \( A \), for which (III) holds. Therefore we can only take \( A_q = L_{ia} & R_{jb} \), and the left side of the equation, as well as the right, will be 1. But the condition that for all \( i, j, a, b, q \) (III) holds cannot be satisfied, because when \( i \neq k \) or \( j \neq l \), the probability \( P(L_{ia} & R_{jb}|L_i & R_j & L_{kc} & R_{ld}) \) is not well-defined, for the event on which we conditionalize is impossible.\(^4\) So finally we see that van Fraassen’s remark can serve as an indirect suggestion that the interpretation (IIIa) is more appropriate. But as we have seen, this interpretation does not suffice to derive (DET).

We can easily convince ourselves that in the framework proposed by van Fraassen no “surface” state can serve as a parameter \( A \) satisfying the requirement (IIIb), no matter whether we introduce the condition (V) or not. It is so because finally only two results of experiments are available to our empirical tests. Hence the parameter \( A \) must remain “hidden”, not in the sense that it cannot ever be discovered, but that it cannot be expressed entirely in terms of “actual settings” or “actual results”. It must unavoidably contain the information of “what would be, had we chosen different settings of the apparatus”. This role could be played, for instance, by the propositions of the form \( L_{ia}^* (R_{jb}^*) \), as introduced earlier. But remember that we explicitly stated (in footnote 3) that some of these propositions should remain neither true nor false. So what we can know at best is the probability of obtaining certain results in certain would-be settings, and that information is insufficient to predict these results with certainty.

5. The Common Cause Reconsidered

Now we must confront the problem of justification of the stronger condition (IIIb), necessary for deriving (DET) and therefore for deriving the Bell inequality. In order to do this we must return to the initial definition of the common cause, proposed after Reichenbach. And here a curious thing becomes apparent. It appears that one element of Reichenbach’s original definition is missing from van Fraassen’s own formulation. Reichenbach namely claims that in order to call an event \( C \) the common cause for \( A \) and \( B \), not only \( C \) but also \( \neg C \) should screen off correlated events \( A \) and \( B \), which amounts to the fourth requirement (cf. Reichenbach 1956, p. 159):

\[
(4) \ P(A & B|\neg C) = P(A|\neg C) P(B|\neg C)
\]

\(^4\) This point can be made even more clear when we recall that the stronger interpretation (IIIb) together with the unquestioned (I) leads to the thesis of determinism (DET). But it is obvious that the hidden variable of the form \( A_q = L_{ia} & R_{jb} \) cannot fulfill the requirement of determinism, for even actual results in one setting cannot determine the “would-be” results in different settings. Hence we see that van Fraassen’s example cannot possibly fulfill the stronger condition (IIIb).
Intuitively, this new requirement should be clear. If we had only condition (3) fulfilled, then in all cases in which \( C \) did not occur, the correlation between \( A \) and \( B \) would be still unexplained. Only when the correlation also disappears for the cases when not-\( C \) occurs, we can claim that \( C \) acts as a causal explanation for the correlation in question, and therefore as the common cause for \( A \) and \( B \). Strangely enough, van Fraassen omitted this important part of Reichenbach’s explication of the common cause. It is even more surprising when we realize that in van Fraassen’s other writings the definition in question was put in the right form (cf. van Fraassen 1980, p. 28 and p. 121). Maybe it was just an accident, a typographic error or something like that. We cannot be sure. But one fact is significant. We have already noted that just right after defining his notion of the common cause with the condition (4) missing, van Fraassen remarks that the informal condition (1) is no less important than the other two, for without it the definition could be satisfied by simply putting \( C = A \& B \). True, it is straightforward that in this case \( P(A \& B|A \& B) = 1 \) and \( P(A|A \& B) = 1, P(B|A \& B) = 1 \). But if we take (4) into considerations, we can easily verify that such a \( C \) does not work. Obviously \( P(A \& B|\text{not-(}A \& B\text{)}) \) equals zero, but \( P(A|\text{not-(}A \& B\text{)}) \) and \( P(B|\text{not-(}A \& B\text{)}) \) need not to be zero. Unless the occurrence of \( A \) implies the occurrence of \( B \) or vice versa, these probabilities are greater than nil. So van Fraassen’s remark was right only with respect to his condition put on the notion of the common cause, and not to Reichenbach’s.5

There is one more interpretative problem connected with the analyzed definition, namely the role and purpose of condition (2). It is significant that van Fraassen never uses it in his entire argument. In order to properly understand the meaning of the condition (2), we must again refer to Reichenbach. It appears that the role of (2) is entirely auxiliary—it merely served Reichenbach in the derivation of the statistical dependence between \( A \) and \( B \) (the condition \( P(A \& B) \neq P(A)P(B) \)). Reichenbach’s idea was to include into his conditions (1)–(4) not only the information that \( C \) explains the correlation between \( A \) and \( B \), but also the information that such a correlation really took place. Moreover, in van Fraassen’s application to the EPR situation condition (2) is in fact violated. It follows immediately from the thesis (DET), which implies that for some \( Aq \) and \( Lia \) the probability \( P(Lia|Li \& Aq) = 0 \). But in such a case this probability cannot be greater than \( P(Lia|Li \& \text{not-Aq}) \).6 So, it would be better to simply give up on condition (2).

We shall now formulate a version of Reichenbach’s definition of the common cause suitable for the purpose of analyzing EPR-like correlations. I propose the following definition:

The statistical correlation between events \( A \) and \( B \) is causally explained iff there exists a family of mutually exclusive and jointly exhaustive events \( \{C_i\}_{i=1,n} \) such that:

5 On the other hand we can easily check that all of Reichenbach’s formal conditions imposed on the notion of the common cause are satisfied by choosing \( C = A \) (or, symmetrically, \( B \)). Obviously \( P(A \& B|A) = P(A) \), and because \( P(A|A) = 1 \), the condition (3) is met. The requirement (4) is trivially satisfied, because \( P(A \& B|\text{not-}A) = P(A|\text{not-}A) = 0 \). So we see that each of the events \( A \) and \( B \) separately can be formally treated as a common cause, but none of them occurs in the absolute past of \( A \) and \( B \), and therefore cannot be considered a proper common cause.

6 I owe this observation to my student, Dariusz Syska.
(a) for every i \( P(A \& B|C_i) = P(A|C_i) P(B|C_i) \)

(b) every \( C_i \) occurs in the absolute past of both \( A \) and \( B \).

The above definition is a straightforward generalization of Reichenbach’s definition, which can be seen as a limit case for \( n = 2 \). The idea behind this condition is quite simple. If we have two statistically dependent events \( A \) and \( B \), then in order to account for this dependence we can either postulate a direct causal link between them, or a common cause in their past. If it is the second case, then we should be able to divide the entire probability space into mutually exclusive events in such a way that for every such event the correlation would disappear. Now the correlation between \( A \) and \( B \) would arise only as a “side effect” of the probability distribution over the family \( C_i \) and the causal correlation between \( C \) and \( A \) and between \( C \) and \( B \) separately. In other words, producing the event \( A \) would per se have no impact whatsoever on the probability of the occurrence of \( B \), as there is no direct “ontological” link between \( A \) and \( B \). The link between \( A \) and \( B \) is merely “epistemological”: in the absence of exact knowledge about any \( C_i \) occurring, we can infer from the occurrence of \( A \) that certain probabilities of \( C_i \) would change, which in turn changes our expectations concerning the likelihood of the occurrence of \( B \).

Now let us apply this general definition to the case of EPR correlations. In principle what is required by the above definition is that for every choice of settings \( L_i \& R_j \) there be a separate parameter \( A_{ij} \) with values \( q_{ij} \) and such that in the ensemble of particles having the same value of this parameter all correlations between experimental outcomes in this setting disappear. So in other words our key assumption would have the following form:

\[
\text{IIIc. For every } i, j \text{ there is a parameter } A_{ij} \text{ such that for all its values } q_{ij} \]
\[
P(L_{ia} & R_{jb}|L_i & R_j & A_{ij}q_{ij}) = P(L_{ia}|L_i & R_j & A_{ij}q_{ij}) P(R_{jb}|L_i & R_j & A_{ij}q_{ij})
\]

From (IIIc) we can easily prove, as in the original van Fraassen’s derivation, that values of every parameter \( A_{ii} \) uniquely determine results of measurements of \( L_i \) as well as \( R_i \):

\[
\text{(DET')} P(L_{ia}|L_i & A_{ii}q_{ii}) = 0 \text{ or } 1
\]
\[
P(R_{ia}|R_i & A_{ii}q_{ii}) = 0 \text{ or } 1
\]

and, moreover, that \( P(L_{ia}|L_i & A_{ii}q_{ii}) = 0 \) iff \( P(R_{ia}|R_i & A_{ii}q_{ii}) = 1 \). From (DET’) it follows that for every parameter \( A_{ii} \) we can divide the set \( I_i \) of all its possible values \( q_{ii} \) into two subsets \( I^1_i \) and \( I^0_i \), meaning that \( q_{ii} \in I^1_i \) iff \( P(L_{i1}|L_i & A_{ii}q_{ii}) = 1 \) (and hence iff \( P(R_{i0}|R_i & A_{ii}q_{ii}) = 1 \)) and \( q_{ii} \in I^0_i \) iff \( P(L_{i0}|L_i & A_{ii}q_{ii}) = 1 \). Of course we must assume that all parameters \( A_{ii} \) had possessed determinate values even before the setting for a measurement was chosen. Therefore we can finally

\[7\] It is interesting that in order to obtain (DET’) it suffices to use a weaker assumption than (IIIc), namely that only perfect correlations between results of experiments in the same settings (\( L_i \) and \( R_i \)) have their common causes.
define a new parameter \( A \) without indices, taking values \( q \) from one of the eight sets \( I(a_{ij},a_{kl},a_{m}) \) given in the following way: \( q \in I(a_{ij},a_{kl},a_{m}) \) iff \( q_{11} \in I_{11}^{a_{ij}} \) & \( q_{22} \in I_{22}^{a_{kl}} \) & \( q_{33} \in I_{33}^{a_{m}} \). We can see that the parameter \( A \) uniquely determines results of all measurements in all available settings, and that when value \( q \) of this parameter lies in \( I(a_{ij},a_{kl},a_{m}) \), the results would be \( L1a_{1}, L2a_{2}, L3a_{3} \) and \( R1(1-a_{1}), R2(1-a_{2}), R3(1-a_{3}) \).

The last step on van Fraassen’s way to the Bell’s inequality is the following. He defines the joint probabilities well known from the original derivation of the Bell-Wigner inequality:

\[
\begin{align*}
    P(1, 2) &= P(L1 & R2 | L1 \& R2) \\
    P(2, 3) &= P(L2 & R3 | L2 \& R3) \\
    P(1, 3) &= P(L1 & R3 | L1 \& R3)
\end{align*}
\]

Next he claims that the following equation holds:

\[
P(1, 2) = \sum_{q \in I} P(L11 & R21 | L1 & R2 & \cdot \cdot \cdot Aq) P(Aq),
\]

and analogously for the other two probabilities.

But this step is a little bit too fast. It looks like an instance of a well-known theorem from probability theory, that we have already employed in our previous discussions, but in fact it should have a slightly different form, namely:

\[
P(1, 2) = \sum_{q \in I} P(L11 & R21 | L1 & R2 & Aq) P(Aq | L1 \& R2)
\]

Now we can appeal to premise (V) and replace the conditional probability \( P(Aq | L1 \& R2) \) with the probability \( P(Aq) \).\(^8\) Thus this premise is necessary to obtain the result needed. Next, using the partition we can easily see that all the summation can be reduced to the subset \( I_{1,0,1} \cup I_{1,0,0} \). So finally we have:

\[
\begin{align*}
P(1, 2) &= \sum_{q \in I_{1,0,1}} P(L11 & R21 | L1 & R2 & Aq) P(Aq) + \\
&+ \sum_{q \in I_{1,0,0}} P(L11 & R21 | L1 & R2 & Aq) P(Aq),
\end{align*}
\]

which reduces to:

\(^8\) Note that to obtain (V) it is not enough to assume that every parameter \( A_{ij} \) is statistically independent of the choice of settings: \( P(A_{ij}) = P(A_{ij}|Lj \& Rk) \). It is so because in general even if events \( A_{ij} \) and \( A_{kl} \) are separately independent of a certain event \( B \), still their conjunction \( A_{ij} \& A_{kl} \) can be statistically correlated with \( B \).
\[ P(1, 2) = P(C101) + P(C100). \]

In the same way we can prove that the following equations hold:

\[ P(2, 3) = P(C110) + P(C010), \]
\[ P(1, 3) = P(C101) + P(C100), \]

and this leads inevitably to the Bell-Wigner inequality:

\[ P(1, 2) + P(2, 3) \geq P(1, 3). \]

Without taking out the conditionalizations of \( Aq \) on experimental settings this move would be impossible, for the summands in the equations for \( P(1, 2) \), \( P(1, 3) \) and \( P(2, 3) \) would be different. So we have finally achieved van Fraassen’s goal, but the road leading to it was more winding than it seemed before. We had to correct the central assumption of the argument, as well as straighten up some lines of reasoning, not to mention the clarification of some obscure remarks.

It may be worth mentioning that in (Hawthorne, Silberstein 1995) another way of deriving the Bell inequality from the assumption of the common cause is proposed. The crucial difference between van Fraassen’s approach and the approach used in this article is that in the latter there is no presupposition made about the existence of perfect correlations, and therefore the determinist thesis (DET) is not derivable. The authors accept assumptions similar to van Fraassen’s hidden locality (IV) and hidden autonomy (V), as well as the assumption of screening off the correlations by the hidden parameter, and some theorems of the classical probability calculus. Using these they derive a slightly different version of the Bell inequality, which is known to be violated in quantum mechanics. So it appears that the hypothesis of the common cause doesn’t help in the explanation of quantum statistical correlation even in the case when there are no perfect, deterministic correlations between results of distant measurements, and when no deterministic hidden-variables are invoked.

But let us return to van Fraassen’s original argument. Is he right in claiming that the experience demands abandoning causal models of explanation? As we have shown, the crucial point of the entire reasoning is the requirement put on the notion of the common cause. If we use the obvious generalization of Reichenbach’s original notion, then the conclusion is inevitable. On the other hand, if we took van Fraassen’s own reformulation for granted (which is, as I suggested, not justified), Bell’s inequality wouldn’t follow. So one can think of finding an intermediate notion of the common cause, weaker than the original Reichenbach’s concept but stronger than van Fraassen’s unfortunate formulation seems to suggest. At the end of this survey we shall briefly explore this possibility.
6. Common Causes in An Indeterministic World

Wesley Salmon once suggested in a slightly different context that Reichenbach’s definition should be altered when applied to truly indeterministic events (Salmon 1978). In order to do this, he proposed to replace the condition (3) in the definition of the common cause by the condition:

\[ (3') P(A \& B | C) > P(A | C) \cdot P(B | C), \]

leaving all the other conditions (1), (2), (4) intact. The situation in which (3’) obtains Salmon calls “an interactive fork”. Salmon illustrates his notion of the interactive fork with the help of a quantum process—namely Compton scattering. He notices that in an indeterministic world we cannot expect that the common cause of a statistical correlation will screen off correlated events. For example, when a photon with initial energy \( E \) impinges on an electron, there is a certain probability that the electron will “bounce off” with energy \( E_1 \), and in this case the photon will have energy \( E_2 = E - E_1 \). So we have a strong correlation between the fact that after collision the electron has energy \( E_1 \) and the fact the photon has energy \( E_2 \)—when one fact occurs, the other must occur with probability 1. How to explain this correlation? In a deterministic world (like in the case of billiard balls) we could say that there are certain parameters characterizing the incident photon, which determine exactly the energy of the electron hit as well as the final energy of the photon, so these parameters could serve as a common cause in Reichenbach’s sense. But according to quantum mechanics the exact value of energy of the electron after the collision is not determined by the quantum characteristics of the incident photon, and therefore there cannot be a common cause fulfilling Reichenbach’s conditions. On the other hand, it is natural to explain causally the correlation in question by the fact that initial energy of the photon was \( E \), in the sense that when this condition is present, the rules of quantum mechanics predict that the correlation will take place. So Salmon concludes that the Reichenbach’s requirement should be weakened to the form given above.

A quick look at EPR correlations reveal that we cannot apply Salmon’s notion directly to our case, because we have here an instance of anticorrelation rather than correlation, so the probability of the joint occurrence of \( L_{ia} \) and \( R_{ia} \) is zero, which therefore cannot be greater than any product of probabilities. But we can easily adjust Salmon’s notion to our case assuming that in the case of anticorrelations the requirement (3) should be replaced by

\[ (3'') P(A \& B | C) < P(A | C) \cdot P(B | C). \]

In response to this proposal van Fraassen argues that this kind of a “causal explanation” involves a regress, for we still should demand an explanation, why in the presence of a factor \( C \) the occurrence of \( A \) (\( B \)) changes the probability of the occurrence of \( B \) (\( A \)) (cf. van Fraassen 1980, p. 30). I think that van Fraassen is right, but we can defend Salmon’s view by pointing out that revealing the existence of the \( C \) fulfilling \( (3'') \) together with (1) and (4) is still better than nothing in terms of explanation. Here is what I mean.
When we are faced with a newly discovered case of statistical correlation between certain events \( A \) and \( B \), the first hypothesis which comes to mind is that these events must be somehow causally correlated, so we could in principle lower or raise the probability of one of them by producing the other. Or in other words, we can hypothesize that each case of occurrence of one of these events makes some change in physical conditions in the vicinity of the other event, and therefore changes its probability. Reichenbach’s notion of the common cause gives us another possible explanation: \( A \) and \( B \) may be in fact not connected at all, and their alleged correlation can be entirely due to some event \( C \) occurring in their common past. But Salmon suggests an intermediate solution: sometimes we can divide further all cases of statistical correlation into cases in which this correlation disappears (not-\( C \)), and cases, in which it is still there (\( C \)). If it is so, we cannot claim, as in the first case, that the occurrence of \( A \) alone lowers or raises the probability of \( B \)—it is \( A \) together with \( C \) which makes \( B \) more or less probable. So it can be argued that the following procedure constitutes a kind of explanation of the correlation: the procedure in which we narrow initial conditions of this correlation as closely as possible in such a way that in all situations when these conditions don’t appear, there is no correlation. In other words, in this kind of explanation we try to pin down conditions “responsible” for the occurrence of the correlation, setting aside all irrelevant factors. It is obvious that we can find in science numerous examples of resorting to such an explanation.

However it is not clear whether this kind of “causal explanation” can help much in the case which van Fraassen presented us with. Namely, the best candidate for Salmon’s common cause of EPR-correlations seems to be the fact that all particles \( L \) and \( R \) were produced in a special quantum state, called the singlet state. Only in this case the correlation would appear. Unfortunately, all of van Fraassen’s considerations are implicitly restricted to particles which already are in the singlet state. When defining the probability space, van Fraassen assumed that we are dealing only with such particles (this was expressed in the assumption (I) concerning perfect correlations). In terms of experimental procedure it means that we have prepared particles in the desired state, and when we compare the result for a particle \( L \) with the result for a particle \( R \), we must be sure that these particles form one pair (this is assured by setting a time interval between detection of the particle \( L \) and the particle \( R \), which must be less than a certain limit). So the way in which he presented the problem prevents us from using further the quantum state of two particles \( L \) and \( R \) as the Salmon common cause for the correlation. Of course, in general it is imaginable that a situation takes place, in which we are confronted with some empirically testable correlations, for which we don’t have any explanation of this sort—just bare empirical fact of correlation, no preparation, no theoretical knowledge. In such a case it would be desirable to find out if there is a factor such that in the absence of it the correlation doesn’t come out. But this simply does not apply to our case.

But perhaps there is a way to further narrow the initial characteristics of particles in the EPR state so that the correlation would occur only within this limit. Let’s see what can be made out of this. We should assume that for every combination \( L_i & R_j \) there is a certain parameter \( A_{ij} \) such that when its actual value lies in a
particular set of values \( I \), the correlations between \( L_{ia} \) and \( R_{jb} \) will disappear, but when this value lies outside this set, the correlation will be still present. And now it is not difficult to notice that this requirement amounts exactly to the weaker condition (IIIa), when the parameter \( A \) can be defined simply as the n-tuple of the parameters \( A_{ij} \). Therefore, it leads only to the weaker thesis of determinism (DET’), which states that every possible result of a measurement is determined by some values of \( A \), and not necessarily the same for each setting. There will be still values of \( A \) such that when \( A \) takes them, some results of measurements will be undetermined, and some correlations will remain unexplained. In consequence, there is no threat of deriving the Bell inequality, but this is at the cost of not solving entirely the mystery of the particular cases of correlation. And apart from this, it is highly improbable that such a parameter \( A \) could be found at all. Nothing indicates even a slightest chance of the existence of a condition narrowing the set of correlated experimental results in the EPR case. So finally it seems that there is no escape from van Fraassen’s no-go result.

References


MAUDLIN, T.: 1994, Quantum Non-locality and Relativity, Basil Blackwell

READHEAD, M.: 1987, Incompleteness, Nonlocality and Realism


