Independent Evidence about a Common Cause

Elliott Sober

Philosophy of Science, Vol. 56, No. 2 (Jun., 1989), 275-287.

Stable URL:
http://links.jstor.org/sici?sici=0031-8248%28198906%2956%3A2%3C275%3AIEAACC%3E2.0.CO%3B2-U

Philosophy of Science is currently published by The University of Chicago Press.

INDEPENDENT EVIDENCE ABOUT A COMMON CAUSE*

ELLIOTT SOBER†

Department of Philosophy
University of Wisconsin, Madison

To infer the state of a cause from the states of its effects, independent lines of evidence are preferable to dependent ones. This familiar idea is here investigated, the goal being to identify its presuppositions. Connections are drawn with Reichenbach's (1956) and Salmon's (1984) discussions of the principle of the common cause.

In the *Philosophical Investigations,* Wittgenstein alludes to a man who is doubtful about the reliability of a story he reads in the newspaper, so he buys another copy of the same newspaper to double check. The joke here is meant to point to a principle: *two observations are better than one only to the extent that the observations are independent.* If different copies of a newspaper always say the same thing, looking at a second copy is simply wasted effort.

Past experience has shown that it is worthwhile to provide intuitive epistemological principles with an explicit and somewhat formal representation. Intuition may suggest that black ravens always confirm the hypothesis that all ravens are black, but a formal representation of this idea shows that it is overstated. The evidential significance of a "positive instance" depends on empirical background assumptions (Good 1967; Rosenkrantz 1977). The same may be said of the idea that only nomological generalizations are confirmed by their instances and the idea that generalizations are confirmed only if predictions about the next instance are too (Sober 1988a). It is the purpose of this paper to discover what assumptions underlie the idea that the confirmatory value of multiple observations turns on their independence.

I

The problem raised by Wittgenstein's story involves inferring the state of a cause from the state of one or more of its effects. A newspaper reports which baseball team won a particular game. The report is the

*Received March 1987; revised June 1987.

†My thanks to Ellery Eells, Malcolm Forster, and the anonymous referees of this journal for comments on an earlier draft.

Philosophy of Science, 56 (1989) pp. 275–287.

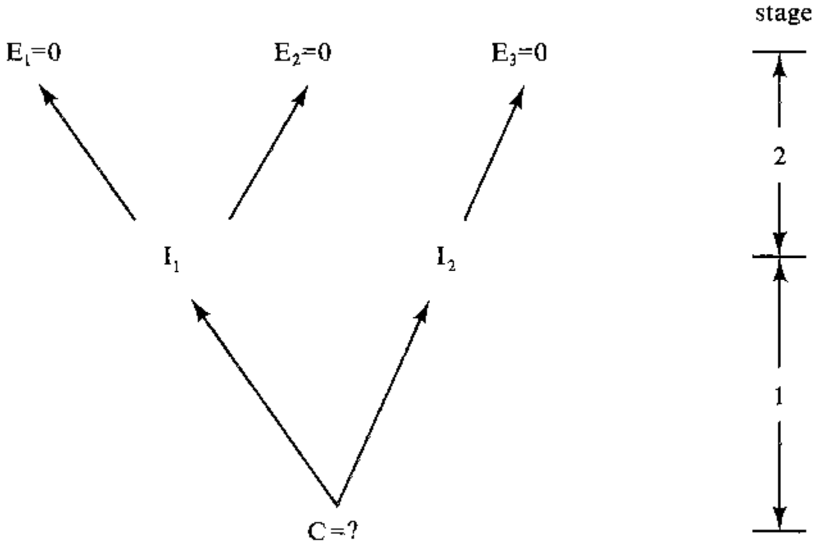

Figure 1

effect and the game is the cause. If the newspaper says that one team was victorious, is this evidence that that team indeed won or that the other won instead? Having reached some verdict about the evidential meaning of a single observation, we then need to judge the significance of reports found in two copies of the same newspaper, as well as the significance of what two independently derived observations say.

So our comparison is three-fold. We must compare the import of a single observation, the import of a pair of nonindependent observations, and the import of a pair of independent observations. To provide the newspaper example with some generality, I will represent the problem in terms of Figure 1 above. Each of the events C, I_1, I_2, E_1, E_2, E_3, comes in two states, which will be coded as "0" or "1". We wish to discover how a 0 state found in one or more of the ultimate effects (E_1, E_2, E_3) bears on an inference about whether the cause C was in state 0 or 1.

Notice that E_1 and E_2 share a cause that is not a cause of E_3. This is the intermediate event I_1. In the newspaper example, E_1 and E_2 might be copies of the same newspaper, whereas E_3 would be an independently derived report on, say, the radio. The two newspaper copies trace back to an arrangement of type that was set for the presses, but the radio announcement does not include this event in its causal history.

The two states 0 and 1 that each event in this branching structure might occupy may be interpreted however one pleases, with a 0 or 1 state in one event having some totally different significance from a 0 or 1 state in another. For example, we might think of the two states possible at C as indicating which team actually won the baseball game, the two states at the intermediate link I_1 as electronic states of a teletype machine in a newsroom, the two states possible at E_1 and E_2 as ink marks on copies of a newspaper, and the two states at E_3 as possible acoustical events transmitted over the radio.

The three-fold comparison of the evidential import of a single observation, two nonindependent observations, and two independent observations involves comparing the answers to three questions: What does the fact that $E_1 = 0$ tell us about the state of C? What does the fact that $E_1 = 0$ and $E_2 = 0$ tell us about the state of C? What does the fact that $E_1 = 0$ and $E_3 = 0$ (or that $E_2 = 0$ and $E_3 = 0$) tell us about the state of C?

I stipulate the following probabilities, so that the three ultimate effects will be assumed "similar" to each other, save for facts about their relative independence or nonindependence. This will allow us to isolate the evidential significance of their relative independence, quite apart from other ways in which they may differ:

$$\Pr(I_1 = 0 / C = 0) = \Pr(I_2 = 0 / C = 0) = s_1$$

$$\Pr(I_1 = 0 / C = 1) = \Pr(I_2 = 0 / C = 1) = e_1$$

$$\Pr(E_1 = 0 / I_1 = 0) = \Pr(E_2 = 0 / I_1 = 0)$$

$$= \Pr(E_3 = 0 / I_2 = 0) = s_2$$

$$\Pr(E_1 = 0 / I_1 = 1) = \Pr(E_2 = 0 / I_1 = 1)$$

$$= \Pr(E_3 = 0 / I_2 = 1) = e_2$$

The notation is meant to suggest that the processes leading from the root to the tips of this tree come in two stages, indicated by the subscripts "1" and "2". Notice that comparable events on different branches (within the same stage) are assumed to have the same probability. An "s" codes for "stasis" (the end of a stage being in the "same" state as the beginning), whereas an "e" indicates "evolution" (the end of the stage being in a state "different" from the one at the beginning). I also assume that more proximal causes screen off more distal causes from their effects, and that causes also screen off noncauses.[1]

[1]For example, $\Pr(E_1 = 0 / I_1 = 0) = \Pr(E_1 = 0 / I_1 = 0 \ \& \ C = 0)$ and $\Pr(E_1 = 0 / I_1 = 0) = \Pr(E_1 = 0 / I_1 = 0 \ \& \ I_2 = 0)$.

The above probabilities do not permit us to derive posterior probabilities for hypotheses specifying the state of event C. The reason is that no prior probabilities have been provided. However, no such calculation is needed, since the question before us can be answered by considering the *likelihoods* of hypotheses about the state of C, relative to different possible observations of the effects E_1, E_2, and E_3. That is, our question about the significance of different possible observations can be answered by seeing what probabilities those observations have in the light of hypotheses about the state of C.

So I begin with six likelihoods. We want to find the probabilities that each hypothesis about the state of C confers on each of three possible observations, corresponding to the cases of a single observation ($E_1 = 0$), two nonindependent observations ($E_1 = 0$ and $E_2 = 0$), and two independent observations ($E_1 = 0$ and $E_3 = 0$):

$$\Pr(E_1 = 0/C = 0) = s_1s_2 + (1 - s_1)e_2 \quad (1)$$

$$\Pr(E_1 = 0 \text{ and } E_2 = 0 / C = 0) = s_1s_2^2 + (1 - s_1)e_2^2 \quad (2)$$

$$\Pr(E_1 = 0 \text{ and } E_3 = 0 / C = 0) = [s_1s_2 + (1 - s_1)e_2]^2 \quad (3)$$

$$\Pr(E_1 = 0 / C = 1) = e_1s_2 + (1 - e_1)e_2 \quad (4)$$

$$\Pr(E_1 = 0 \text{ and } E_2 = 0 / C = 1) = e_1s_2^2 + (1 - e_1)e_2^2 \quad (5)$$

$$\Pr(E_1 = 0 \text{ and } E_3 = 0 / C = 1) = [e_1s_2 + (1 - e_1)e_2]^2. \quad (6)$$

Notice that the algebraic representation alone does not say whether the single observation favors $C = 0$ over $C = 1$. Further assumptions about the processes involved are required to settle whether (1) > (4). The familiar world we inhabit is one in which newspaper reports about baseball games are usually to be taken at their word. But we can easily imagine a situation in which such reports have the opposite significance or no significance at all.

I will represent the significance of the three possible observations in terms of three ratios. If the single observation favors the hypothesis that $C = 0$, this means that the ratio of (1) to (4) is greater than unity. Similarly, the significance of the two nonindependent observations will be represented by the ratio of (2) to (5) and the import of the two independent observations by the ratio of (3) to (6). The question posed by Wittgenstein's example can then be investigated by seeing what relationships obtain among these three ratios:

$$R_{1,4} = [s_1s_2 + (1 - s_1)e_2]/[e_1s_2 + (1 - e_1)e_2]$$

$$R_{2,5} = [s_1s_2^2 + (1 - s_1)e_2^2]/[e_1s_2^2 + (1 - e_1)e_2^2]$$

$$R_{3,6} = [s_1s_2 + (1 - s_1)e_2]^2/[e_1s_2 + (1 - e_1)e_2]^2.$$

The first ratio, recall, represents the evidential significance of a single observation, the second, the bearing of a pair of nonindependent observations, and the third, the import of two independent observations.

As noted before, no prior probabilities for $C = 0$ and $C = 1$ are specified in the model here investigated. But suppose these were included, with an eye to determining how different possible observations affect the overall plausibility (=posterior probability) of hypotheses about C. If so, a comparison of ratios of posterior probabilities would be in order, and that comparison would turn just on the likelihood ratios just described.

The first thing to notice about these ratios is that they will be identical, if all component probabilities take on extreme values of 0 or 1. Suppose, for example, that all the s terms are 1 and the e terms are 0. This makes all three ratios infinite; the three observations all favor the hypothesis that $C = 0$ to the same (maximal) degree. In this deterministic setting it is pointless to look at a second copy of the same newspaper to double check the testimony of the first. But it is equally pointless to turn on the radio for an "independent" check. Epistemological differences among the three sources of evidence collapse in this extreme case.

So let us assume that some component probabilities do not take extreme values. The most straightforward conclusion now concerns the relationship of $R_{1,4}$ and $R_{3,6}$. Here we compare the evidential significance of a single observation with that of a pair of independent observations. Notice that the latter is just the square of the former. This means that $R_{3,6}$ exceeds $R_{1,4}$ if and only if $R_{1,4}$ is greater than unity.

So if the zero state of an effect is evidence favoring the hypothesis that C was in the zero state, this implies that two independent effects in that zero state confirm that hypothesis about the cause more than a single one does. It is important to notice that this result is conditional. No absolute claim is advanced here that when both the newspaper and the radio independently say that team 0 was victorious, that this supports the hypothesis that team 0 was the winner more than either observation does by itself. All we can say is that *if* each report points to a certain interpretation when taken by itself, then the two point to that same interpretation with added weight.

Having reached a verdict on the comparable worth of a single observation and two independent observations, we now consider how a single observation compares to the testimony of two nonindependent observations. We therefore compare $R_{1,4}$ and $R_{2,5}$. Notice first that these ratios will be identical, if the transition probabilities pertaining to the second stage of the process (s_2 and e_2) take on extreme values. Regardless of what is true of the probabilities concerning the first stage, there is no point in looking at a second copy of the same newspaper, if the two copies will necessarily say the same thing.

However, if the second stage probabilities are intermediate, $R_{1,4} < R_{2,5}$ if and only if

$$s_2(s_1 - e_1) > e_2(s_1 - e_1). \tag{7}$$

This also happens to be the very condition under which $R_{1,4} > 1$. So even with highly dependent joint effects, two matching observations are better than one. If a single effect's being in state zero favors the hypothesis that C was in state 0, then two dependent effects' both being in state zero favors that hypothesis all the more (providing, of course, that the second stage of the process is not deterministic).

We turn, finally, to comparing the testimony of two dependent observations with the testimony of two independent observations—to the relationship of $R_{2,5}$ and $R_{3,6}$. The former is less than the latter precisely when

$$e_1 s_1 s_2{}^2[s_1 - e_1] > (1 - e_1)(1 - s_1)e_2{}^2[s_1 - e_1]. \tag{8}$$

We have already seen that, if probabilities are intermediate, the condition that settles the evidential meaning of a single observation also shows that two dependent observations are better than one. This is inequality (7). Condition (8), which governs whether the two independent observations are better evidence than the two dependent observations, is logically independent of condition (7). The first array of values makes (7) true and (8) false, whereas the second makes (7) false and (8) true:

$$s_1 = 0.4 \quad e_1 = 0.3 \quad s_2{}^2 = 0.9 \quad e_2{}^2 = 0.8$$

$$s_1 = 0.7 \quad e_1 = 0.6 \quad s_2{}^2 = 0.8 \quad e_2{}^2 = 0.9.$$

It is worth noting that (7) is true precisely when:

$s_i > e_i$ for each i, or $s_i < e_i$ for each i.

Although this condition does not suffice to verify (8), it suffices if supplemented with a "symmetry" condition asserting that the probability of change within a branch is the same regardless of whether it involves going from 0 to 1 or from 1 to 0:

$$e_i = (1 - s_i), \text{ for each } i.[2] \tag{*}$$

I summarize these arguments about the three ratios as follows: relatively weak assumptions suffice to show why two observations are better than one, regardless of whether the two are dependent or independent.

[2]Note that (*) also says that the probability of stasis is the same, whether it involves going from 0 to 0 (s_i) or from 1 to 1 ($1 - e_i$).

But more needs to be assumed to show that independent observations offer better confirmation than dependent ones.

II

In comparing $R_{2,5}$ and $R_{3,6}$, I compared how two dependent effects and two independent effects provide evidence about the state of their common cause. Although it is not inevitably true that $R_{2,5} < R_{3,6}$, there is a range of cases in which the states of the independent effects (E_1, E_3) provide more evidence than the states of the dependent effects (E_1, E_2). It is interesting to note how this fact relates to Reichenbach's (1956) and Salmon's (1984) discussion of an idea they term the principle of the common cause.

Their idea, roughly, is that correlated events provide evidence for postulating a common cause. If there were a correlation between the inscriptions found in two copies of a newspaper, this would be evidence that they trace back to a common cause. If their reports were uncorrelated, on the other hand, there would be no need to postulate a common cause. So the Reichenbach/Salmon principle *seems* to suggest that correlated (nonindependent) events provide better evidence about common causes than uncorrelated (independent) ones do. Yet, when we compared dependent and independent effects, we found circumstances in which *independent* effects provide more powerful evidence. Does this mean that the Reichenbach/Salmon principle and the ideas developed here are in conflict?

To see that this conflict is only apparent, we must distinguish unconditional from conditional independence. Events E_1 and E_3 are independent, conditional on the state of C. That is, $\Pr(E_1 = i \text{ and } E_3 = j/C = k) = \Pr(E_1 = i/C = k) \Pr(E_3 = j/C = k)$, for all $i, j, k = 0, 1$. On the other hand, if we do not conditionalize on the state of C, the states of E_1, and E_3 can be expected (under a general circumstance to be specified presently) to be correlated. It is just this unconditioned correlation that the Reichenbach/Salmon principle says is evidence for postulating a common cause, one that renders the joint effects conditionally probabilistically independent.

Although I have my doubts about the correctness of the principle of the common cause (Sober 1987, 1988b), the present discussion does not contradict it. For we must recognize that the Reichenbach/Salmon principle pertains to postulating the existence of a common cause, not, in the first instance, to deciding what state that common cause occupies. If a newspaper report and a radio broadcast are correlated, perhaps we should infer that they trace back to a common cause—say, to a baseball game they both describe. But if we want to know what actually happened in that baseball game, we proceed to a separate inference problem. It is one thing to infer the existence of a common cause, quite another to say what that common cause was like, on the assumption that it exists.[3]

The inference problem explored here *assumes* that the three effects trace back to a common cause C. This is not inferred from the evidence, but is a condition imposed at the outset. Given that C exists and bears the causal connections described in Figure 1, we wish to know whether it was in state 0 or in state 1. To answer this question, there are assumptions that may lead us to prefer information about the (conditionally) independent effects of that common cause, and other assumptions that ground the opposite preference. Although the following principle cannot claim a universal validity, it does have a range of conditions in which it is correct: strong unconditional correlation between two events is good if you want to infer the existence of a common cause, but strong conditional correlation is bad if you want to infer the state of that common cause, having already assumed that it exists.

This contrast may be fleshed out by using covariance as a measure of the strength of association between pairs of events. The covariance of X and Y is defined as $\Pr(X \text{ and } Y) - \Pr(X)\Pr(Y)$. Independent events have a covariance of zero. Although E_1 and E_3 have zero covariance if we conditionalize on the state of C, we will calculate covariances without assuming that this state is known. To do this, we need a prior probability for the states of the cause. Letting $\Pr(C = 0) = a$, we obtain

$$\begin{aligned}\mathrm{Cov}(E_1, E_2) &= a[s_1s_2^2 + (1 - s_1)e_2^2] \\ &\quad + (1 - a)[e_1s_2^2 + (1 - e_1)e_2^2] - \Pr(E_1)\Pr(E_2).\\ \mathrm{Cov}(E_1, E_3) &= a[s_1s_2 + (1 - s_1)e_2]^2 \\ &\quad + (1 - a)[e_1s_2 + (1 - e_1)e_2]^2 - \Pr(E_1)\Pr(E_3).\end{aligned}$$

It turns out that $\mathrm{Cov}(E_1, E_2) > \mathrm{Cov}\ (E_1, E_3) > 0$, if

all probabilities are intermediate and $\quad s_i \neq e_i\ (i = 1, 2)$. (**)

Recall from (7) that the inequality in (**) is essential if a single observation is to be evidentially relevant to the state of the cause. So (**) describes a highly general circumstance; when it is satisfied, more closely related effects will have a higher covariance than more distantly related ones, with both sorts of events showing a positive association.

Let us suppose, then, that E_1, E_2, and E_3 are such that two of them

[3]In applying the principle of the common cause to examples, Reichenbach (1956) and Salmon (1984) often treat postulating a common cause and inferring the state of that cause interchangeably. See, for example, Salmon's (1984, pp. 213–221) discussion of Perrins' estimation of the value of Avogadro's number and the analysis of the murder mystery example that follows it.

have a common cause not shared by the third. In this case, if (**) is true, we may infer, from observing the frequencies of 0 states in the three events, that $\mathrm{Cov}(E_1,E_2) > \mathrm{Cov}(E_1,E_3) > 0$, and from this that the causal relationship of the three events is given by the tree structure represented in Figure 1. However, the assumptions licencing this inference of the existence of I_1 do not settle whether the observational pair $E_1 = 0$ and $E_3 = 0$ provides better evidence concerning the state of C than the pair $E_1 = 0$ and $E_2 = 0$ does; that is, (**) does not imply either (8) or its negation.

An example, due to Salmon (1984), may make this contrast more graphic. Two students are suspected of each copying a list of random numbers (in binary notation, suppose) from a book in the library. Presumably, the strength of the correlation between the lists is a measure of how plausible it is to think that they plagiarized from the same source. But now, instead of inferring that the students plagiarized, let us assume that they did so. We now wish to infer what the list in the library from which they copied actually was like. Which of the following two arrangements would we prefer the students to have used? The two students could have copied from a *shared* intermediary, who had gone to the library and copied from the book. Or the two students could have copied from *separate* intermediaries, each of whom independently went to the book and copied from it. These scenarios correspond to events E_1 and E_2 and events E_1 and E_3. Plausible assumptions, given by (**), may indicate that there would be a smaller covariance of the two resulting lists if the students had copied from separate intermediaries. Yet, in terms of recovering what the library book actually said, this arrangement may make for a stronger inference.

Notice that the inference of a common cause by appeal to the fact of positive covariance is based on the observed frequencies displayed by *kinds* of events. Though the two students each have lists of numbers in which 0's and 1's occur about half the time, it turns out that if the first student has a 0 in the *i*th position, the second student probably does too. We survey the *entirety* of both lists of random numbers to infer the existence of a common cause.

However, if we want to know whether the book in the library has a 0 or a 1 at the *i*th position, it is entirely irrelevant (we may assume) what the students' lists include at positions other than *i*. To infer the state of this part of the common cause, we focus not on the frequencies of *kinds* of events, but on two *token* events: the two students both have a 0 at the *i*th position of their lists. Given this fact about the matching of token events, rather than the correlation of types of events, we then proceed to ask how the state of the common cause could be inferred.[4]

[4] It is possible to interpret the discussion of ratios $R_{1,4}$, $R_{2,5}$, and $R_{3,6}$ as pertaining to the

The issues considered here find an application in the context of phylogenetic inference. One may think of the E_i as species, all of which trace back to an ancestor C. One wishes to infer what traits the ancestor exhibited from information about the traits of its descendents. For example, imagine that E_1 and E_2 are humans and chimps, whereas E_3 is a species of kangaroo. C, we may imagine, is a hypothetical mammalian ancestor. If humans and chimps resemble each other in some respect, this may offer some evidence that the distant ancestor had that trait as well. But it is presumably far better evidence about the state of this ancestor to find the resemblance between two less closely related species—say, between humans and kangaroos. Here the evolutionary inference favors putting greater weight on the matching of traits found in more independent lines.

This inference, plausible though it may be, is not without its presuppositions. It is not enough to assume that the phylogenetic tree has the shape described in Figure 1. Additionally, one must make assumptions about the transition probabilities attaching to its separate branches. The assumption of uniform rates of evolution within lineages is not enough either.[5] Even within the context of that assumption, (8) states what must be true for the resemblance of more distantly related taxa to provide greater weight of evidence as to the ancestral form.

Although (7) and (*) suffice to vouchsafe this conclusion, I do not say that they are "assumptions" of the intuitive evolutionary judgment. Assumptions are the conditions that must be true, for an inference to make sense; I have argued that (7) and (*) are jointly sufficient, not that they are necessary. Nor is (8), properly speaking, a general assumption behind the idea that independent evidence is superior to nonindependent evidence. The reason is that (8) embodies nonessential features of the comparative problem, preeminently the assumption of sameness of transition probabilities on simultaneous branches.

This last failure of generality is easily correctable. We could assign each branch in Figure 1 its own uniquely labeled transition probabilities and then set forth an inequality of the required sort. Since that representation would no longer exploit the assumption of uniform rates, the resulting inequality would have a better claim on being termed an assumption of the independence principle than the special case investigated above. Its nontriviality, like that of (8), would point once again to a general

---

evidential significance of kinds of observations, rather than to tokens. My point here is that the Reichenbach/Salmon principle takes correlation of event types as evidence, whereas my discussion has focused on the matching of event tokens.

[5]Here I use "uniform" to mean that all branches at a given time have the same transition probabilities; this does not imply that branches at different times have the same transition probabilities. Uniformity does not mean constancy (Sneath and Sokal 1973, p. 321).

conclusion of this paper: no matter what the details are of the particular process model under investigation, substantive assumptions must be made, if judgments about the relative weight of dependent and independent evidence are to be advanced.

III

The above discussion has focused on how one should interpret a zero state found in one or more of the three events displayed on the tips of the tree. This is a different matter from designing an experiment whose results are not foreseeable beforehand. That is, I have described the evidential significance of finding that E_1 and E_3 are both in state zero, but this is not the same as saying whether it is worth examining these two events to discover what states they occupy. This latter problem about the design of experiments would have to consider the significance of all four possible outcomes—of finding that the two events were in states 1/1, 1/0, 0/1, or 0/0. To assess the value of an experiment, one must consider all its possible results; but in describing the evidential significance of a single outcome—say, the observation of 0 states in E_1 and E_2—one need not consider the results that could have occurred, but did not. This is an instance of Hacking's (1965, 1971) distinction between *before trial betting* and *after trial evaluation*.

Insofar as deciding to carry out an experiment is a decision about action, considerations besides the evidential significance of possible outcomes are relevant. For example, time and money might lead a scientist to run one experiment instead of another, even if the results of the second would be expected to provide more decisive evidence than the results of the first. However, imagine that these "non-epistemic" values are not germane; suppose the only concern in designing an experiment is to obtain the most telling evidence possible about the state of the cause. How should this decision problem proceed?

It is not to be doubted that looking at all three effects is superior to looking at any subset thereof. In just the same way, the results of the previous sections nowhere contradict a principle of total evidence. If the states of all three effects are known, one should consult these, not some contracted data set made of just one or two observations.[6] But just as we considered before how the significance of two observations would be affected by the discovery that they are dependent or independent, we now may consider how we should choose between examining the two dependent and the two independent effects.

[6]Malcolm Forster has pointed out to me that the likelihood ratio, relative to the observation $E_1 = 0$ and $E_2 = 0$ and $E_3 = 0$, exceeds the pair-wise and singleton ratios defined before.

If we knew the values of the branch probabilities displayed in the figure, we could calculate probabilities for each possible pair-wise observation. For each possible observation, we also could calculate a likelihood ratio, which measures how profoundly that observation would favor one hypothesis about the cause's state over the other. In this case, the likelihood ratio is a measure of utility—it indicates how well the evidence discriminates between the two competing hypotheses. The task now is to calculate the expected utility of the experiment from these two sets of information.

The pertinent measure will be an "expected ratio of likelihood", where the likelihood ratio associated with each pair-wise observation is weighted by that observation's probability of occurring. To implement this idea, we would first have to normalize the ratios, so that values less than unity can contribute to the expectation to the same degree as values that are greater. This might be achieved by taking inverses of ratios less than unity.[7] Having calculated the expected utility of each experiment as a weighting over the utilities of its four possible results, we then would choose to run the experiment with the greater expected utility.

The calculation just described requires that we know the values of the probabilistic parameters associated with the branches in the figure. However, there are circumstances in which an experimental design can be chosen with less than full information about these values. For example, suppose I stipulate that two matching effects are evidence that the cause was in the same state as the effects, but that a pair of nonmatching effects provide no information at all about the cause. Can I then choose between examining the dependent and examining the independent effects?

If assumption (**) holds true, the dependent effects have a higher probability of matching than do the independent ones. If (8) is correct, a matching between independent observations provides more evidence than a match between dependent effects. Let us assume both (**) and (8). This means that if I examine the two dependent effects, there is a higher probability that I will obtain relevant information—a match. However, if I examine the independent effects but manage nonetheless to obtain a match, that information will be weightier than a match obtained on dependent effects.

To decide what to do, I must choose a decision rule. I could be conservative; by examining the dependent effects, I maximize my chance of obtaining relevant information, however paltry. Or I could gamble; by examining the independent effects, I run a greater risk of obtaining useless observations, but a match between independent effects, should I ob-

[7]I leave open the possibility that other problems will need to be solved to obtain a reasonable measure of an experiment's expected utility.

serve one, would be worth more than a matching of dependent effects. I do not need to know the point values of the probabilistic parameters to make this choice; rather, I need to choose between maximin and maximax strategies.

In discussing the evidential significance of different observations before, I nowhere had to have numerical values for the probabilistic parameters associated with the branches in the figure; inequalities were established that obtain independently of what those exact values turn out to be. Likewise, the interpretation of an observed match did not demand the choice of a decision rule. Yet when the problem is designing an experiment, rather than interpreting one of its possible results, these additional sorts of information turn out to be essential. Before you run an experiment, you generally know less about the system under study than you do once the experiment has produced results. It is for this reason that a problem arising earlier can require more information for its solution than one that arises subsequently.

REFERENCES

Good, I. (1967), "The White Shoe Is a Red Herring", *British Journal for the Philosophy of Science 17*: 322.

Hacking, I. (1965), *The Logic of Statistical Inference*. Cambridge: Cambridge University Press.

———. (1971), "Jacques Bernouilli's Art of Conjecturing", *British Journal for the Philosophy of Science 22*: 209–229.

Reichenbach, H. (1956), *The Direction of Time*. Berkeley: University of California Press.

Rosenkrantz, R. (1977), *Inference, Method, and Decision*. Dordrecht: Reidel.

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

Sneath, P., and Sokal, R. (1973), *Principles of Numerical Taxonomy*, 2nd ed. San Francisco: W. H. Freeman.

Sober, E. (1987), "Explanation and Causation: A Review of Wesley Salmon's *Scientific Explanation and the Causal Structure of the World*", *British Journal for the Philosophy of Science 38*: 243–257.

———. (1988a), "Confirmation and Lawlikeness", *Philosophical Review 97*: 93–98.

———. (1988b), *Reconstructing the Past: Parsimony, Evolution, and Inference*. Cambridge, Mass.: Bradford/MIT Press.