Causal Laws and Effective Strategies

NANCY CARTWRIGHT

STANFORD UNIVERSITY

INTRODUCTION

There are at least two kinds of laws of nature: laws of association and causal laws.[1] Laws of association are the faimilar laws with which philosophers usually deal. These laws tell how often two qualities or quantities are co-associated. They may be either deterministic—the association is universal—or probabilistic. The equations of physics are a good example: *whenever* the force on a classical particle of mass m is f the acceleration is f/m. Laws of association may be time indexed, as in the probabilistic laws of Mendelian genetics, but apart from the asymmetries imposed by time indexing, these laws are causally neutral. They tell how often two qualities co-occur; but they provide no account of what makes things happen.

Causal laws, by contrast, have the word "cause"—or some causal surrogate, right in them. Smoking causes lung cancer; perspiration attracts wood ticks; or,—for an example from physics, force causes change in motion: to quote Einstein and Infeld ([5]: 9), "The action of an external force changes the velocity. . . such a force either increases or decreases the velocity according to whether it acts in the direction of motion or in the opposite direction."

Bertrand Russell [9] argued that laws of association are all the laws there are, and that causal principles cannot be derived from the causally symmetric laws of association. I shall here argue in support of Russell's second claim, but against the first. Causal principles cannot be reduced to laws of association; but they cannot be done away with.

The argument in support of causal laws relies on some facts about strategies. They are illustrated in a letter which I recently received from TIAA-CREF, a company which provides insurance for college teachers. The letter begins:

> It simply wouldn't be true to say,
> "Nancy L. D. Cartwright. . . if you own a TIAA life insurance policy you'll live longer."
> But it is a fact, nonetheless, that persons insured by TIAA do enjoy longer lifetimes, on the average, than persons insured by commercial insurance companies that serve the general public.

I will take as a starting point for the argument facts like those reported by the TIAA letter: *it wouldn't be true that* buying a TIAA policy would be an effective strategy for lengthening one's life. TIAA may, of course, be mistaken; it could after all be true. What is important is that their claim is, as they suppose, the kind of claim which is either true or false. There is a pre-utility sense of goodness of strategy; and what is, and what is not, a good strategy in this pre-utility sense is an objective fact. Consider a second example. Building the canal in Nicaragua, the French discovered that spraying oil on the swamps is a good strategy for stopping the spread of malaria, whereas burying contaminated blankets is useless. What they discovered was true, independent of their theories, of their desire to control malaria, or of the cost of doing so.

The reason for beginning with some uncontroversial examples of effective and ineffective strategies is this: I claim causal laws cannot be done away with, for they are needed to ground the distinction between effective strategies and ineffective ones. If indeed, it *isn't true that* buying a TIAA policy is an effective way to lengthen one's life, but stopping smoking is, the difference between the two depends on the causal laws of our universe, and on nothing weaker. This will be argued in Part II. Part I endorses the first of Russell's claims, that causal laws cannot be reduced to laws of association.

I.1 STATISTICAL ANALYSES OF CAUSATION

I will abbreviate the causal law, *C causes E* by $C \rightarrow E$. Notice that *C* and *E* are to be filled in by general terms, and not names of particulars; for example, *Force causes motion* or *Aspirin relieves headache*. The generic law *C causes E* is not to be understood as a universally quantified law about particulars, even about particular causal facts. It is generically true that aspirin relieves headache even though some particular aspirins fail to

do so. I will try to explain what causal laws assert by giving an account of how causal laws relate on the one hand to statistical laws, and on the other to generic truths about strategies. The first task is not straightforward, for, although causal laws are intimately connected with statistical laws, they cannot be reduced to them.

A primary reason for believing that causal laws cannot be reduced to probabilistic laws is broadly inductive: no attempts so far have been successful. The most notable attempts recently are by the philosophers Patrick Suppes [12] and Wesley Salmon [10] and, in the social sciences by a group of sociologists and econometricians working on causal models, of whom Herbert Simon and Hubert Blalock (cf. Blalock [2]) are good examples.

It is not just that these attempts fail, however, but rather why they fail that is significant. The reason is this. As Suppes urges, a cause ought to increase the frequency of its effect. But this fact may not show up in the probabilities if other causes are at work. Background correlations between the purported cause and other causal factors may conceal the increase in probability which would otherwise appear. A simple example will illustrate.

It is generally supposed that smoking causes heart disease ($S \rightarrow H$). Thus, we may expect that the probability of heart disease on smoking is greater than otherwise. (We can write this as either $P(H/S) > P(H)$, or $P(H/S) > P(H/-S)$, for the two are equivalent.) This expectation is mistaken, however. Even if it is true that smoking causes heart disease, the expected increase in probability will not appear if smoking is correlated with a sufficiently strong preventative, say exercising. (Leaving aside some niceties, we can render *Exercise prevents heart disease* as $X \rightarrow -H$.) To see why this is so, imagine that exercising is more effective at preventing heart disease than smoking at causing it. Then in any population where smoking and exercising are highly enough correlated,[2] it can be true that $P(H/S) = P(H)$, or even $P(H/S) < P(H)$. For the population of smokers also contains a good many exercisers, and when the two are in combination, the exercising tends to dominate.

It is possible, however, to get the increase in conditional probability to reappear. The decrease arises from looking at probabilities which average over both exercisers and non-exercisers. Even though, in the general population, it seems

better to smoke than not, in the population consisting entirely of exercisers, it is worse to smoke. This is also true in the population of non-exercisers. The expected increase in probability occurs—not in the general population—but in both sub-populations.

This example depends on a fact about probabilities known as Simpson's paradox (Simpson [11]), or sometimes as the Cohen-Nagel-Simpson paradox, for it is presented as an exercise in Morris Cohen's and Ernest Nagel's text, *An Introduction to Logic and Scientific Method* [4]. Nagel suspects that he learned about it from G. Yule's *An Introduction to the Theory of Statistics* (1904), which is one of the earliest textbooks written on statistics; and indeed it is discussed at length there. The fact is this: any association—$P(A/B) = P(A)$; $P(A/B) > P(A)$; $P(A/B) < P(A)$—between two variables which holds in a given population can be reversed in the subpopulations by finding a third variable which is correlated with both.

In the smoking—heart disease example, the third factor is a preventative factor for the effect in question. This is just one possibility. Wesley Salmon [10] has proposed different examples to show that a cause need not increase the probability of its effect. His examples also turn on Simpson's paradox, except that in his cases the cause is correlated, not with the presence of a negative factor, but with the absence of an even more positive one.

Salmon considers two pieces of radioactive material, uranium 238 and polonium 214. We are to draw at random one material or the other, and place it in front of a geiger counter for some time. The polonium has a short half-life, so that the probability for some designated large number of clicks is .9; for the long-lived uranium, the probability is .1. In the situation described, where one of the two pieces is drawn at random, the total probability for a large number of clicks is $\frac{1}{2}(.9) + \frac{1}{2}(.1) = .5$. So the conditional probability for the geiger counter to click when the uranium is present is less than the unconditional probability. Nevertheless, when the uranium has been drawn and the geiger counter does register a large number of clicks, it is the uranium that causes them. The uranium decreases the probability of its effect in this case. But this is only because, as Salmon constructs it, whenever the uranium is present, the even more effective polonium is absent.

All the counter examples I know to the claim that causes increase the probability of their effects work in this same way. In all cases, the cause fails to increase the probability of its effects for the same reason: in the situation described the cause is correlated with some other causal factor which dominates in its effects. This suggests that the condition as stated is too simple. A cause must increase the probability of its effects—but only in situations where such correlations are absent.

The most general situations in which a particular factor is not correlated with any other causal factors are situations in which all other causal factors are held fixed, that is, situations which are homogeneous with respect to all other causal factors. In the population where everyone exercises, smoking cannot be correlated with exercising. So, too, in populations where no-one is an exerciser. I hypothesize then that the correct connection between causal laws and laws of association is this:

> *C causes E* if and only if *C* increases the probability of *E* in every situation which is otherwise causally homogeneous with respect to *E*.

Carnap's [3] notion of a state description can be used to pick out the causally homogeneous situations. A complete set of causal factors for *E* is the set of all C_i such that either $C_i \rightarrow +E$ or $C_i \rightarrow -E$. (For short $C_i \rightarrow \pm E$.) Every possible arrangement of the factors from a set which is complete except for *C* picks out a population homogeneous in all causal factors but *C*. Each such arrangement is given by one of the 2^n state descriptions $K_j = \bigwedge \pm C_i$ over the set $\{C_i\}$ (*i* ranging from 1 to n) consisting of all alternative causal factors. These are the only situations in which probabilities tell anything about causal laws. I will refer to them as *test* situations for the law $C \rightarrow E$.

Using this notation the connection between laws of association and causal laws is this:

CC: $C \rightarrow E$ iff $P(E/C.K_j) > P(E/K_j)$ for all state descriptions K_j over the set $\{C_i\}$, where $\{C_i\}$ satisfies

(i) $C_i \in \{C_i\} \Rightarrow C_i \rightarrow \pm E$

(ii) $C \notin \{C_i\}$

(iii) $\forall D\ (D \rightarrow \pm E \Rightarrow D = C \text{ or } D \in \{C_i\})$

(iv) $C_i \in \{C_i\} \Rightarrow \neg(C \rightarrow C_i)$.

Condition (iv) is added to ensure that the state descriptions do not hold fixed any factors in the causal chain from C to E. It will be discussed further in the section after next.

Obviously *CC* does not provide an analysis of the schema $C \rightarrow E$, for exactly the same schema appears on both sides of the equivalence. It does, however, impose mutual constraints, so that given sets of causal and associational laws cannot be arbitrarily conjoined.

Initially, the condition *CC* appears extremely weak. It looks as if, given a set of associational laws, it takes all $n - 1$ causal laws to fix whether the n^{th} obtains or not. This appearance is misleading; some probability measures, in fact, are consistent with only one set of causal laws.[3] These are nice cases, however. In general more than one set of causal laws could be conjoined to a single set of probability laws. Thus two worlds could be identical in all their laws of association, yet differ in their causal laws. Humeans will find this unsatisfactory. Nevertheless, *CC* is, I believe, the strongest connection that can be drawn between causal laws and laws of association.

I.2. TWO ADVANTAGES FOR SCIENTIFIC EXPLANATION

C. G. Hempel's [7] original account of inductive-statistical explanation had two crucial features which have been given up in later accounts, particularly in Salmon's: 1) an explanatory factor must increase the probability of the fact to be explained; 2) what counts as a good explanation is an objective, person-independent matter. Both of these features seem to me to be right. If we use causal laws in explanations, we can keep both these requirements and still admit as good explanations just those cases that are supposed to argue against them.

i) Hempel insisted that an explanatory factor increase the probability of the phenomenon it explains. This is an entirely plausible requirement although there is a kind of explanation for which it is not appropriate. In one sense, to explain a phenomenon is to locate it in a nomic pattern. The aim is to lay out all the laws relevant to the phenomenon; and it is no matter to this aim whether the phenomenon has high or low probability under these laws. Although this seems to be the kind of explanation that Richard Jeffrey describes in "Statistical Explanation vs. Statistical Inference," [8] it is not the kind

of explanation that other of Hempel's critics have in mind. Salmon, for instance, is clearly concerned with causal explanation.[4] Even for causal explanation Salmon thinks the explanatory factor may decrease the probability of the factor to be explained. He supports this with the uranium-plutonium example described above.

What makes the uranium count as a good explanation for the clicks in the geiger counter, however, is not the probabilistic law Salmon cites (P(clicks/uranium) < P(clicks)), but rather the causal law—*Uranium causes radioactivity*. As required, the probability for radioactive decay increases when the cause is present, *for every test situation*. There is a higher level of radioactivity when uranium is added both for situations in which polonium is present, and for situations in which polonium is absent. Salmon sees the probability decreasing because he attends to a population which is not causally homogeneous.

Insisting on increase in probability across all test situations not only lets in the good cases of explanation which Salmon cites; it also rules out some bad explanations that must be admitted by Salmon. For example, consider a case which, so far as the law of association is concerned, is structurally similar to Salmon's uranium example. I consider eradicating the poison oak at the bottom of my garden by spraying it with defoliant. The can of defoliant claims that the spray is 90 per cent effective; that is, the probability of a plant's dying given that is sprayed is .9, and the probability of its surviving is .1. Here, in contrast to the uranium case, only the probable outcome, and not the improbable, is explained by the spraying. One can explain why some plants died by remarking that they were sprayed with a powerful defoliant; but this will not explain why some survive.[5]

The difference is in the causal laws. In the favorable example, it is true both that uranium causes high levels of radioactivity and that uranium causes low levels of radioactivity. This is borne out in the laws of association. Holding fixed other causal factors for a given level of decay, either high or low, it is more probable to get that level if uranium is added than not. This is not so in the unfavorable case. It is true that spraying with defoliant causes death in plants, but it is not true that spraying also causes survival. Holding fixed other causes of death, spraying with my defoliant will increase the probability of a plant's dying; but holding fixed other causes

of survival, spraying with that defoliant will decrease, not increase the chances of a plant's surviving.

ii) All these explanations are explanations by appeal to causal laws.[6] Accounts, like Hempel's or Salmon's or Suppes', which instead explain by appeal to laws of association, are plagued by the reference class problem. All these accounts allow that one factor explains another just in case some privileged statistical relation obtains between them. (For Hempel the probability of the first factor on the second must be high; for Suppes it must be higher than when the second factor is absent; Salmon merely requires that the probabilities be different.) But whether the designated statistical relation obtains or not depends on what reference class one chooses to look in, or on what description one gives to the background situation. Relative to the description that either the uranium or the plutonium is drawn at random, the probability of a large number of clicks is lower when the uranium is present than otherwise. Relative to the description that plutonium and all other radio-active substances are absent, the probability is higher.

Salmon solves this problem by choosing as the privileged description the description assumed in the request for explanation. This makes explanation a subjective matter. Whether the uranium explains the clicks depends on what information the questioner has to hand, or on what descriptions are of interest to him. But the explanation which Hempel aimed to characterize was in no way subjective. What explains what depends on the laws and facts true in our world, and cannot be adjusted by shifting our interest or our focus.

Explanation by causal law satisfies this requirement. Which causal laws are true and which are not is an objective matter. Granted, certain statistical relations must obtain: the cause must increase the probability of its effect. But no reference class problem arises. In how much detail should we describe the situations in which this relation must obtain? We must include all and only the other causally relevant features. What interests we have, or what information we focus on, is irrelevant.

I.3. SOME DETAILS AND SOME DIFFICULTIES

Before carrying on to Part II, some details should be noted and some defects admitted.

a) *Condition (iv)*. Condition (iv) is added to the above characterization to avoid referring to singular causal facts. A test situation for $C \rightarrow E$ is meant to pick out a (hypothetical infinite) population of individuals which are alike in all causal factors for E—except those which on that occasion are caused by C itself. The test situations should not hold fixed factors in the causal chain from C to E. If it did so, the probabilities in the populations where all the necessary intermediate steps occur would be misleadingly high; and where they do not occur, misleadingly low. Condition (iv) is added to except factors caused by C itself from the description of the test situation. Unfortunately it is too strong. For condition (iv) excepts any factor which *may* be caused by C even on those particular occasions when the factor occurs for other reasons. Still, (*iv*) is the best method I can think of for dealing with this problem, short of introducing singular causal facts, and I let it stand for the nonce.

b) *Interactions*. One may ask, "But couldn't it happen that $P(E/C) > P(E)$ in *all* causally fixed circumstances, and still C not be a cause of E?" I don't know. I am unable to imagine convincing examples in which it occurs; but that is hardly an answer. One kind of example, however, is clearly taken account of. That is the problem of spurious correlation (sometimes called "the problem of joint effects"). If two factors E_1 and E_2 are both effects of a third factor C, then it will frequently happen that the probability of the first factor is greater when the second is present than otherwise, over a wide variety of circumstances. Still, we do not want to assert $E_1 \rightarrow E_2$. According to principle *CC*, however, $E_1 \rightarrow E_2$ only if $P(E_1/E_2) > P(E_1)$ both when C obtains, and also when C does not obtain. But the story that E_1 and E_2 are joint effects of C provides no warrant for expecting either of these increases.

One may have a worry in the other direction as well. Must a cause increase the probability of its effect in *every* causally fixed situation? Mightn't it do so in some, but not in all? I think not. Whenever a cause fails to increase the probability of its effect, there must be a reason. Two kinds of reasons seem possible. The first is that the cause may be correlated with other causal factors. This kind of reason is taken account of. The second is that interaction may occur. Two causal factors are interactive if in combination they act like a single causal factor whose effects are different from at least one of the two

acting separately. For example, ingesting an acid poison may cause death; so too may the ingestion of an alkali poison. But ingesting both may have no effect at all on survival.

In this case, it seems, there are three causal truths: (1) ingesting acid without ingesting alkali causes death; (2) ingesting alkali without ingesting acid causes death; and (3) ingesting both alkali and acid does not cause death. All three of these general truths should accord with *CC*.

Treating interactions in this way may seem to trivialize the analysis: anything may count as a cause. Take any factor which behaves sporadically across variation of causal circumstances. May we not count it as a cause by looking at it separately in those situations where the probability increases, and claim it to be in interaction in any case where the probability does not increase? No. There is no guarantee that this can always be done. For interaction is always interaction with some other *causal factor*; and it is not always possible to find some other factor, or conjunction of factors, which obtain just when the probability of E on the factor at issue decreases, and which itself satisfies principle *CC* relative to all other causal factors (See for example "How Some Worlds Could Not be Hume Worlds," note 12). Obviously, considerably more has to be said about interactions, but this fact at least makes it reasonable to hope they can be dealt with adequately, and that the requirement of increase in probability across all causal situations is not too strong.

c) *0, 1 probabilities and threshold effects*. Principle *CC* as it stands does not allow $C \rightarrow E$ if there is even a single arrangement of other factors for which the probability of E is one, independent of whether C occurs or not. So *CC* should be amended to read:

$$C \rightarrow E \text{ iff } (\forall_j)\ \{P(E/C.K_j) > P(E/K_j) \text{ or } P(E/K_j) = 1 = P(E/C.K_j)\} \text{ and } (\exists_j)\ \{P(E/K_j) \neq 1\}$$

It is a consequence of the second conjunct that something which occurs universally can be the consequent of no causal laws. The alternative is to let anything count as the cause of a universal fact.

There is also no natural way to deal with threshold effects, if there are any. If the probability of some phenomenon can be raised just so high, and no higher, the treatment as it stands allows no genuine causes for it.

d) *Time and causation. CC* makes no mention of time. The properties may be time indexed—taking aspirins at t causes relief at $t+\Delta t$ but the ordering of the indices plays no part in the condition. Time ordering is often introduced in statistical analyses of causation to guarantee the requisite assymetries. Some for example, take increase in conditional probability as their basis. But the causal arrow is assymetric, whereas increase in conditional probability is symmetric: P(E/C)> P(E) iff P(C/E)> P(C). This problem does not arise for *CC*, because the set of alternative causal factors for *E* will be different from the set of alternative causal factors for *C*. I take it to be an advantage that my account leaves open the question of backwards causation. I doubt that we shall ever find compelling examples of it, but if there were a case in which a later factor increased the probability of an earlier one in all test situations, it might well be best to count it a cause.

II.1. PROBABILITIES IN DECISION THEORY

Standard versions of decision theory require two kinds of information. 1) How desirable are various combinations of goals and strategies; and 2) how effective are various strategies for obtaining particular goals. The first is a question of utilities, which I will not discuss. The second is a matter of effectiveness; it is generally rendered as a question about probabilities. We need to know what may roughly be characterized as "the probability that the goal will result if the strategy is followed." It is customary to measure effectiveness by the conditional probability. Following this custom, we could define

!S is an *effective strategy* for G iff P(G/S) > P(G).

I have here used the volative mood marker ! introduced by H. P. Grice,[7] to be read "let it be the case that." I shall refer to *S* as *the strategy state*. For example, if we want to know whether the defoliant is effective for killing poison oak, the relevant strategy state is *a poison oak plant is sprayed with defoliant*. On the above characterization, the defoliant is effective just in case the probability of a plant's dying, given that it has been sprayed, is greater than the probability of its dying given that it has not been sprayed. Under this characterization, the

distinction between effective and ineffective strategies depends entirely on what laws of association obtain.

But the conditional probability will not serve in this way, a fact which has been urged by Allan Gibbard, and William Harper [6]. Harper and Gibbard point out that the increase in conditional probability may be spurious, and that spurious correlations are no grounds for action. Their own examples are somewhat complex because they specifically address a doctrine of Richard Jeffrey's not immediately to the point here. We can illustrate with the TIAA case already introduced. The probability of long life given that one has a TIAA policy is higher than otherwise. Nevertheless, as the letter says, it would be a poor strategy to buy TIAA in order to increase one's life expectancy.

The problem of spurious correlation in decision theory leads naturally to the introduction of counterfactuals. We are not, the argument goes, interested in how many people have long lives among peoples insured by TIAA. but rather in the probability that one *would have* a long life if one *were* insured with TIAA. Apt as this suggestion is, it requires us to evaluate the probability of counterfactuals, for which we have only the beginnings of a semantics (via the device of measures over possible worlds)[8] and no methodology, much less an account of why the methodology is suited to the semantics. It would be preferable to have a measure of effectiveness that only required probabilities over events, which can be tested in the actual world in the standard ways. This is what I shall propose.

The Gibbard and Harper example, an example of spurious correlation due to a joint cause, is a special case of a general problem. We saw that the conditional probability will not serve as a mark of causation in situations where the putative cause is correlated with other causal factors. Exactly the same problem arises for effectiveness. The conditional probability is not a good measure of effectiveness in any populations where the strategy state is correlated with other factors causally relevant to the goal state, for whatever reason the correlation obtains. Increase in conditional probability is no mark of effectiveness in situations which are causally heterogeneous. It is necessary, therefore, to make the same restrictions about test situations in dealing with strategies that we made in dealing with causes:

!S is an *effective strategy* for obtaining G in situation L iff $P(G/S.K_L) > P(G/K_L)$

Here K_L is the state description true in L, taken over the complete set $\{C_i\}$ of causal factors for G, barring S. L may not fix a unique state description, however. For example L may be the situation I am in when I decide whether to smoke or not, and, at the time of the decision, it is not determined whether I will be an exerciser. In that case, we should compare not the actual values $P(G/S.K_L)$and $P(G/K_L)$, but rather their expected values:

SC: ! S is an *effective strategy* for obtaining G in L iff $\sum_j P(G/S.K_j)P(K_j) > \sum_j P(G/K_j)\ P(K_j)$, where j ranges over all K_j consistent with L.[9]

This formula for computing the effectiveness of strategies has several desired features: (1) It is a function of the probability measure, P, given by the laws of association in the actual world; and hence calculable by standard methods of statistical inference. (2) It reduces to the conditional probability in cases where it ought. (3) It restores a natural connection between causes and strategies.

(1) *SC* avoids probabilities over counterfactuals. Implications of the arguments presented here for constructing a semantics for probabilities for counterfactuals will be pointed out in section II.3.

(2) Troubles for the conditional probability arise in cases, like the TIAA example, in which there is a correlation between the proposed strategy and (other) causal factors for the goal in question. When such correlations are absent, the conditional probability should serve. This follows immediately: when there are no correlations between S and other causal factors, $P(K_j/S) = P(K_j)$; so the left hand side of *SC* reduces to P(G/S)in the situation L and the right hand side to P(G) in L.

(3) There is a natural connection between causes and strategies that should be maintained: if one wants to obtain a goal, it is a good (in the pre-utility sense of good) strategy to introduce a cause for that goal. So long as one holds both the simple view that increase in conditional probability is a sure mark of causation and the view that conditional probabilities are the right measure of effectiveness, the connection is

straightforward. The arguments in Part I against the simple view of causation break this connection. But *SC* re-establishes it, for it is easy to see from the combination of *CC* and *SC* that if $X \rightarrow G$ is true, then $!X$ will be an effective strategy for G in any situation.

II.2. CAUSAL LAWS AND EFFECTIVE STRATEGIES

Although *SC* joins causes and strategies, it is not this connection which argues for the objectivity of *sui generis* causal laws. As we have just seen, one could maintain the connection between causes and strategies, and still hope to eliminate causal laws by using simple conditional probability to treat both ideas. The reason causal laws are needed in characterizing effectiveness, is that they pick out the right properties to condition on. The K_j which are required to characterize effective strategies must range over *all* and *only* the causal factors for G.

It is easy to see, from the examples of Part I, why the K_j must include *all* the causal factors. If any are left out, cases like the smoking-heart disease example may arise. If exercising is not among the factors which K_j fixes, the conditional probability of heart disease on smoking may be less than otherwise in K_j, and smoking will wrongly appear as an effective strategy for preventing heart disease.

It is equally important that the K_j not include too much. $\{K_j\}$ partitions the space of possible situations. To partition too finely it as bad as not to partition finely enough: Partitioning on an irrelevancy can make a genuine cause look irrelevant, or make an irrelevant factor look like a cause. Earlier discussion of Simpson's paradox shows that this is structurally possible. Any association between two factors C and E can be reversed by finding a third factor which is correlated in the right way with both. When the third factor is a causal factor, the smaller classes are the right ones to use for judging causal relations between C and E. In these, whatever effects the third factor has on E are held fixed in comparing the effects of C versus those of $\neg C$. But when the third factor is causally irrelevant to E—that is, when it *has* no effects on E—there is no reason for it to be held fixed, and holding it fixed gives wrong judgments both about causes and about strategies.

I will illustrate from a real life case.[10] The graduate school at Berkeley was accused of discriminating against women in

their admission policies, thus raising the question "*Does being a woman cause one to be rejected at Berkeley?*" The accusation appeared to be borne out in the probabilities: the probability of rejection was much higher for men than for women. Bickel, Hammel, and O'Connel [1] looked at the data more carefully, however, and discovered that this was no longer true if they partitioned by department. In a majority of the 85 departments, the probability of admission for women was just about the same as for men, and in some even higher for women than for men. This is a paradigm of Simpson's paradox. Bickel, Hammel and O'Connell accounted for the paradoxical reversal of associations by pointing out that women tended to apply to departments with high rejection rates, so that department by department women were admitted in about the same ratios as men, but across the whole university, considerably fewer women, by proportion, are admitted.

This analysis seems to exonerate Berkeley from the charge of discrimination. But only because of the choice of partitioning variable. If, by contrast, the authors had pointed out that the associtions reversed themselves when the applicants were partitioned according to their roller skating ability that would count as no defense.[11] Why is this so?

The difference between the two situations lies in our antecedent causal knowledge. We know that applying to a popular department (one with considerably more applicants than positions) is just the kind of thing that causes rejection. But, without a good deal more detail, we are not prepared to accept the principle that being a good roller skater causes a person to be rejected by the Berkeley graduate school, and we make further causal judgments accordingly. If the increased probability for rejection among women disappears when a causal variable is held fixed, the hypothesis of discrimination in admissions is given up; but not if it disappears only when some causally irrelevant variable is held fixed.

The Berkeley example illustrates the general point: only partitions by causally relevant variables count in evaluating causal laws. If changes in probability under causally irrelevant partitions mattered, almost any true causal law could be defeated by finding, somewhere, some third variable that correlates in the right ways to reverse the required association between cause and effect.

II.3. ALTERNATIVE ACCOUNTS WHICH EMPLOY "TRUE PROBABILITIES" OR COUNTERFACTUALS

One may object: Once all causally relevant factors have been fixed, there is no harm to finer partitioning by causally irrelevant factors. For, contrary to what is claimed in the remarks about roller skating and admission rates, further partitioning will not change the probabilities. There is a difference between true probabilities and observed relative frequencies. Granted, it is likely that one can always find some third, irrelevant, variable which, on the basis of estimates from finite data, appears to be correlated with both the cause and effect in just the ways required for Simpson's paradox. We are concerned here, however, not with finite frequencies, or estimates from them; but rather with true probabilities. You misread the true probabilities from the finite data, and think that correlations exist where they do not.

For this objection to succeed, an explication is required of the idea of a true probability and this explication must make plausible the claim that partitions by what are pre-analytically regarded as non-causal factors do not result in different probabilities. It is not enough to urge the general point that the best estimate often differs from the true probability; there must in addition be reason to think that that is happening in every case where too-fine partitioning seems to generate implausible causal hypotheses. This is not an easy task, for often the correlations one would want to classify as "false" are empirically indistinguishable from others that ought to be classified "true." The misleading, or "false", correlations sometimes pass statistical tests of any degree of stringency we are willing to accept as a general requirement for inferring probabilities from finite data. They will often, for example, be stable both across time and across randomly selected samples.

That the task is difficult does not rule it out; and one may well have independent philosophic motivations for distinguishing true from accidental or "false" correlations. If so, the arguments here, will help provide adequacy conditions for a satisfactory account. "True" probabilities must be tailored to do the same job in making inferences about causes and strategies that are done here by holding fixed all and only other causal factors.

Similar remarks apply to counterfactual analyses. One popular kind of counterfactual analysis would have it that

!S is effective strategy for G in L iff Prob($S\Box\!\!\rightarrow G/L$) > Prob ($\neg S\Box\!\!\rightarrow G/L$) [12]

The counterfactual and the causal law approach will agree, only if

A: Prob($\alpha \Box\!\!\rightarrow G/X$) = $P(G/\alpha.K_X)$

where K_X is the maximal *causal* description (barring α) consistent with X. Assuming the arguments here are right, condition A provides an adequacy criterion for any satisfactory semantics of counterfactuals and probabilities.

III. CONCLUSION

The quantity $P(E/C.K_j)$, which appears in both the causal condition of Part I and in the measure of effectiveness from Part II, is called by statisticians the partial conditional probability of E on C, holding K_j fixed; and it is used in ways similar to the ways I have used it here. It forms the foundation for regression analyses of causation and it is applied by both Suppes and Salmon to treat the problem of joint effects. In decision theory, the formula *SC* is structurally identical to one proposed by Brian Skyrms in his deft solution to Newcomb's paradox; and elaborated further in his forthcoming book *Causal Necessity*.[13] What is especially significant about the partial conditional probabilities which appear here, is the fact that these hold fixed all and only causal factors.

The choice of partition, $\{K_j\}$, is the critical feature of the measure of effectiveness proposed in *SC*. This is both a) what makes the formula work in cases where the simple conditional probability fails; and b) what makes it necessary to admit causal laws if you wish to sort good strategies from bad. The way you partition is crucial. For in general you get different results from *SC* if you partition in different ways. (Consider two different partitions for the same space, $K_1 \ldots\ldots, K_n$ and $I_1, \ldots, I_s$, which cross grain each other—the K_i are mutually disjoint and exhaustive, and so are the I_j. Then it is easy to produce a measure over the field ($\pm G$, $\pm C$, $\pm K_i$, $\pm I_j$) Such that

$$\sum_{j=1}^{n} P(G/C.K_j)P(K_j) \neq \sum_{j=1}^{n} P(G/C.I_j)P(I_j)\ .)$$

What partition is employed is, therefore, essential to whether a strategy appears effective or not. The right partition—the one that judges strategies to be effective or ineffective in accord with what is objectively true—is determined by what the causal laws are. Partitions by other factors will give other results; and, if you do not admit causal laws there is no general procedure for picking out the right factors. The objectivity of strategies requires the objectivity of causal laws.

REFERENCES

[1] Bickel, Peter J., Hammel, Eugene A., and O'Connell, J. William. "Sex Bias in Graduate Admissions: Data from Berkeley," in Fairley, William B. and Mosteller, Frederick, *Statistics and Public Policy* (Reading, Mass: Addison-Wesley, 1977).

[2] Blalock, H.M., Jr. *Causal Models in the Social Sciences* (Chicago: Adline-Atherton, 1971).

[3] Carnap, Rudolf. *The Continuum of Inductive Methods* (Chicago: Univ. of Chicago Press, 1952).

[4] Cohen, Morris R. and Nagel, Ernest. *An Introduction to Logic and Scientific Method* (New York: Harcourt, Brace, Co., 1934).

[5] Einstein, Albert and Infeld, Leopold. *The Evolution of Physics* (Cambridge: Cambridge University Press, 1971).

[6] Gibbard, Allan and Harper, William. "Counterfactuals and Two Kinds of Expected Utility." Discussion Paper No. 194. Center for Mathematical Studies in Economics and Management Science. Northwestern University, January, 1976.

[7] Hempel, C. G. *Aspects of Scientific Explanation* (New York: Free Press, 1965).

[8] Jeffrey, Richard C. "Statistical Explanation vs. Statistical Inference," in Salmon, Wesley, *Statistical Explanation and Statistical Relevance* (Pittsburgh: Pittsburgh University Press, 1971).

[9] Russell, Bertrand. "On the Notion of Cause," *Proceedings of the Aristotlean Society*, 13(1912-13), 1-26.

[10] Salmon, Wesley. "Statistical Explanation" in Salmon, Wesley, ed., *Statistical Explanation and Statistical Relevance* (Pittsburgh: Pittsburgh University Press, 1977).

[11] Simpson, E. H. "The Interpretation of Interaction in Contingency Tables," *Journal of the Royal Statistical Society, Ser. B, 13, No. 2(1951), p. 238-241.*

[12] Suppes, Patrick. A Probabilistic Theory of Causality," *Acta Philosophica Fennica*, Fasc. XXIV, 1970.

NOTES

[1]This research was primarily conducted during a Mellon Post-doctoral grant at the University of Pittsburgh; and I am particularly grateful to the History and Philosophy of Science department there for their help and support.

[2]Throughout, "A and B are correlated" will mean $P(A/B) \neq P(A)$.

[3]See, Cartwright, Nancy, "How Some Worlds Could Not Be Hume Worlds," draft 2, Stanford University, unpublished manuscript.

[4]This is explicitly stated in Salmon's later papers (Cf. "Theoretical Explanation" in Korner, S., *Explanation* (Oxford: Basil Blackwell, 1975).), but it is already clear from

the treatment in "Statistical Explanation. . ." that Salmon is concerned with causal explanations, otherwise there is no accounting for his efforts to rule out "spurious" correlations as explanatory.

[5]This example can be made exactly parallel to the uranium-polonium case by imagining a situation in which we choose at random between this defoliant and a considerably weaker one which is only 10% effective.

[6]Iwill not here offer a model of causal explanation, but certain negative theses follow from my theory. Note particularly that falling under a causal law (plus the existence of suitable initial conditions) is neither necessary nor sufficient for explaining a phenomenon.

It is not sufficient because a single phenomenon may be in the domain of various causal laws, and in many cases it will be a legitimate question to ask, "Which of these causal factors actually brought about the effect on this occasion?" This problem is not peculiar to explanation by causal law, however. Both Hempel in his inductive-statistical model and Salmon in the statistical relevance account side step the issue by requiring that a "full" explanation cite all the possibly relevant factors, and not select among them.

Conversely, under the plausible assumption that singular causal statements are transitive, falling under a causal law is not necessary for explanation either. This results from the fact that (as *CC* makes plain) causal laws are not transitive. Hence a phenomenon may be explained by a factor to which it is linked by a sequence of intervening steps, each step falling under a causal law, without there being any causal law which links the explanans itself with the phenomenon to be explained.

[7]Grice, H. P., *Some Aspects of Reason*, The Immanuel Kant Lectures. Stanford University, 1977.

[8]Cf. Harper, William. Proceedings of Western Ontario conference on semantics for conditionals, May, 1978. London, Ontario. Forthcoming.

[9]I first derived this formula by reasoning about experiments. I am especially grateful to David Lewis for pointing out that the original formula was mathematically equivalent to the shorter and more intelligible one presented here.

[10]Roger Rosenkrantz and Persi Diaconis first pointed out to me that the feature of probabilities described here is called "Simpson's paradox," and the reference for this example was supplied by Diaconis.

[11]William Kruskal discusses the problem of choosing a partition for this data briefly in correspondence following the Bickel, Hammel, and O'Connell article, loc. cit., n. 22.

[12]Cf. articles in forthcoming book edited by William Harper referred to in note 8.

[13]Skyrms, Brian "Newcomb Without Tears," University of Illinois, unpublished manuscripts; and *Causal Necessity* (New Haven: Yale University Press, forthcoming).