David Papineau

Pure, Mixed, and Spurious Probabilities and Their Significance for a Reductionist Theory of Causation

1. Introduction

What should we make of such facts as the fact that smoking is probabilistically associated with cancer, or the fact that there is a high correlation between the amount of sun in August in France and the size of the French grape harvest, or the fact that putting a piece of lead under the "1" on a die makes it more likely that the die will land with the "6" face up when thrown? That is, what should we make of facts of the form P(S/C) > P(S/-C)? In particular, I want to ask, what do such correlational facts have to do with causation?

Some philosophers are inclined to say that, give or take a bit, these correlational facts *constitute* causal connections. On this view we have causation as soon as we have positive statistical relevance between two temporally ordered event types, that is, as soon as the probability of the later event given the earlier event is greater than its probability without the earlier event.

I shall call this the "statistical-relevance" view (the S-R view for short). The name most often associated with this view is Wesley Salmon (see in particular Salmon 1970). However, as I shall have occasion to observe later, Salmon's theory of causation has a number of different aspects, and indeed his most recent position is not really a version of the S-R view at all. I shall take Patrick Suppes's (1970) to be the canonical version of the S-R view.

Even the canonical version of the S-R view needs more explanation than has been offered so far. But let me leave this to one side for a moment. For my initial concern is not with the details of the S-R view, but with the whole idea. I want to raise the question of whether we ought to admit indeterministic causation at all.

Contrast the S-R view with the traditional view that causes have to *determine* their effects. On this view, if C doesn't determine E, it doesn't cause it either. Causation isn't just a matter of the cause making the effect more probable, but of its making the effect *certain*.

I want to suggest that this traditional deterministic view is a viable alternative to the S-R view. At first sight this suggestion might seem misguided. Doesn't

modern physics show that indeterminism is a basic feature of the physical world? But this isn't the issue. I don't hold any brief for determinism itself. I accept that determinism is false. The view I want to put forward as an alternative to the S-R view is not determinism, but a deterministic view of *causation*. This allows that certain events aren't determined. It simply insists that, insofar as they aren't determined, then they aren't caused either.

Let me spell the point out. Defenders of the deterministic view of causation will claim that circumstances that don't determine some particular E aren't enough to cause E. But this needn't be because they think that there must be more to be said about the case in question, because they think that we haven't yet succeeded in identifying those further features of the situation which made E inevitable. They can accept that there may have been no such further features. Their claim is simply that if there were no such further features, then E wasn't caused: if the prior circumstances didn't make E inevitable, then they shouldn't be counted as causing E. (Of course if *nothing* were ever determined, then a deterministic view of causation would imply that nothing was ever caused, and that would perhaps be indefensible. But, as we shall see, allowing that some things aren't determined doesn't imply that everything is undetermined. In particular, in the paradigm cases where some E is undetermined, it will still be true that E's having a certain *chance* will be determined by the prior circumstances.)

So the initial issue to be discussed here is the choice between the traditional deterministic and the S-R views of causation. In a previous paper I argued in favor of the traditional view over the S-R view (Papineau 1985a). However, I am no longer as persuaded as I was then that the traditional view holds a definite advantage. I now think that, when all the arguments are taken into account, the result is pretty much a stand-off between the two sides. It doesn't make much difference which view we uphold. Both theories can be made to accommodate all the relevant data, albeit in different ways, and in the end the choice is as much terminological as substantial.

The first aim of this paper is to survey the relevant arguments and show that they lead to a stand-off. This survey will simplify some of the arguments of my 1985a paper, and it will also introduce a number of new considerations. In the course of this survey I shall have occasion to turn to a second issue, an issue which arises independently of whether we take causes to make their effects certain or merely probable. This is the question of whether a full *reduction* of causal facts to facts of correlational dependence is possible. In the last few years a number of philosophers (including Salmon) have become persuaded that causation somehow transcends correlational relationships, that there is some extra cement in causal connections that resists reduction to facts of correlational association. I want to argue, against these writers, that causation can be so reduced.

So I intend to show two things. First, that there is nothing to choose between the S-R and traditional deterministic views. And, second, that whichever we do

choose, we can reduce causal facts to correlational facts—to facts about some events making others more likely, if we prefer the S-R view of causation, or to facts about some events making others certain, if we prefer the deterministic view of causation.

I now think that this second topic, the possibility of a reduction of causation, is a more substantial philosophical issue than the conflict between the S-R view and the traditional deterministic view of causation. The conflict between the S-R and traditional views eventually degenerates into an unimportant trading of preferences. But, even so, it will be worth exploring at some length. For the issues involved, and the distinctions that need to be made, will prove essential to resolving the question of reduction.

2. Some Initial Intuitions

We have conflicting initial intuitions about causation. On the one hand, there is an initial intuition that causation demands determinism. If, on a given occasion, a full specification of the circumstances left it open that E might *not* have occurred, then how can we say those circumstances *caused* E? On other qualitatively identical occasions E sometimes fails to occur. Doesn't this show that those circumstances aren't enough to cause E on their own?

On the other hand, it also seems intuitively plausible that a probabilistic connection between two properties establishes causation. If smoking makes cancer more likely, doesn't this show that smoking *causes* cancer?

Let us examine the deterministic intuition first. This isn't as straightforward as it looks. For it can be argued that the underlying intuition here is not about causation as such, but about *explanation*: it is the intuition that we haven't fully *explained* E if the circumstances we cite don't make E certain.

To translate this into a conclusion about causation itself we need to assume further that causation is inseparable from full explanation, that E couldn't be caused unless it were fully explainable. But this further assumption is contentious. Thus Philip Kitcher thinks that explanation requires certainty, but that causation only requires increased probability (Kitcher 1985, 638). In a slightly different vein, D. H. Mellor thinks that certainty is required for *full* explanation, but holds that explanation, like causation, comes in degrees; and that we have an explanation, albeit not a full one, whenever the circumstances increase the probability of E (Mellor, forthcoming).

And then, of course, there are also philosophers who simply deny the underlying intuition, and who say that if the prior circumstances increased E's probability, then that's as good an explanation as we ever have. (See the papers by Salmon, Richard Jeffrey, and James Greeno in their 1970 publication. Salmon also argues that if a complete description of the circumstances *reduces* the probability of E, then that amounts to a satisfactory explanation too. But we can leave that to one

side for the moment. Whatever other virtues this suggestion might have, it's certainly not supported by any initial intuitions. I shall say something more about such "negative causes" in section 18 below.)

In this paper I want to talk about the objective relationship of causation, not about the anthropocentric idea of explanation. So rather than get bogged down in the connections between the two notions, I shall simply concede the point at issue. In what follows I won't make any further appeal to the intuitive idea that causation requires certainty.

But doesn't this now concede the whole argument to the S-R view? The contrary intuition, that an increase in probability suffices for causation, now seems to have the field to itself. However, I think this intuition is also far less straightforward than it looks. Indeed I think I can show that, despite appearances, this intuition is quite compatible with a deterministic view of causation.

I shall devote the next six sections, 3–8, to explaining this last claim. This will leave the S-R and deterministic views on an equal footing as far as initial intuitions go. I shall then turn, in sections 9–12, to some considerations involving the relation between causation and the rationality of action. These considerations too will leave the issue between the S-R and deterministic views of causation undecided. But they will focus the question of the reducibility of causation. This question will be pursued in sections 13–18. Section 18 will also contain some final comments on the choice between the deterministic and S-R views.

3. Pure and Mixed Probabilities

I shall say that a conditional probability of the form P(C/S) is *mixed* if there exists some Z such that P(C/SZ) $\neq$ P(C/S). If there is no such Z, I shall say the original conditional probability is *pure*.

This distinction is related to the distinction between homogeneous and inhomogeneous partitions of a reference class. If P(C/S) and P(C/-S) are pure, then the partition of the overall reference class into smokers and nonsmokers gives us a homogeneous partition with respect to cancer, in the sense that subdividing the reference class by additional factors Z won't make any further difference to the probability of cancer: given that you are a smoker (or a nonsmoker), nothing else about you makes any difference to the probability of your getting cancer. But if either P(C/S) or P(C/-S) is mixed, then the partition into S and not-S is inhomogeneous, in that further subdivisions, by Z and -Z, will alter the probabilities. The original partition ignored distinctions between S&Z and S&$-$Z (and $-$S&Z and $-$S&$-$Z) which are in fact relevant to the probability of cancer.

Pure probabilities give the single-case *chances* in particular cases. The chance of C in a particular case will be the probability conditional on all the relevant circumstances: and since "relevant circumstances" here means factors whose presence or absence makes a difference to the conditional probability, such chance-

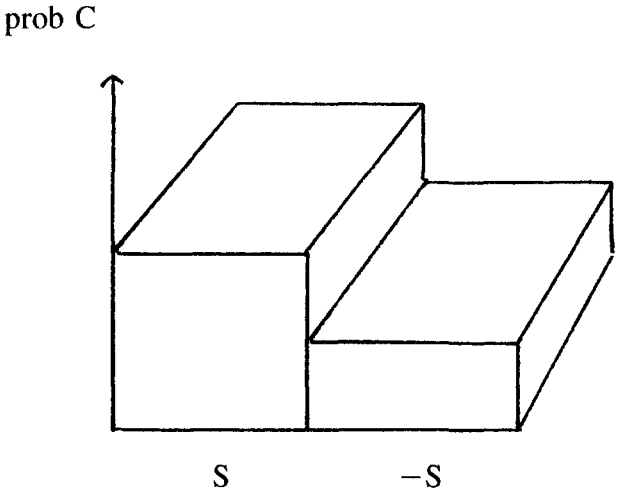

Figure 1. S probabilistically relevant to C.

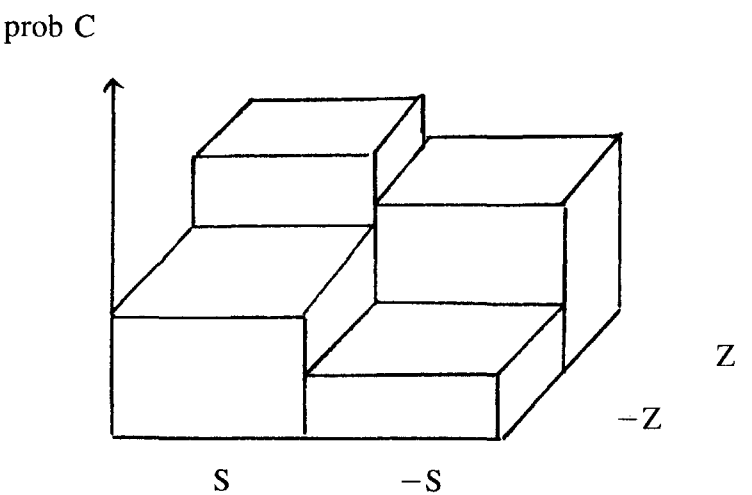

Figure 2. Z mixes P(C/S) and P(C/−S).

giving conditional probabilities will always be pure. Conversely, mixed probabilities are in general no good for giving single-case chances, since further conditionalization on circumstances present in the particular case will generally alter a mixed probability.[1]

I take it that it is uncontroversial that the correlations with which I began are mixed. Surely there are other differences between people, apart from whether

they smoke or not, which affect the probability of their getting cancer. Surely the amount of sunshine in August isn't the only thing that makes a difference to the size of the grape harvest. Surely, even given a loaded die, there is further room for craps to be "a game of skill," in the Runyonesque sense.

What is more, I take it that this mixedness cannot be eliminated simply by appealing to further things we know about. We do know about some of the further things that make a difference to the probability of cancer, such as working in asbestos factories, or living in dirty air, etc. But even when we have partitioned our reference class by such further known factors, there will still no doubt be some further unknown factors making our probabilities mixed: it is scarcely likely that the medical researchers have identified *all* the factors that are probabilistically relevant to cancer. And similarly in the other cases. Even after we have taken into account all the things that the viniculturists and the dice experts know about, there will still be further unknown factors that make a probabilistic difference, and the probabilities that we will be left with will still be mixed. (Note that I am not making the false deterministic assumption that the only real chances are nought and one. I am simply making the uncontroversial point that we are almost certainly ignorant of some of the factors relevant to the real chances.)

4. Screening Off and Spurious Correlations

What do these last remarks, about the mixedness of our original probabilities, imply about their causal status? If those probabilities are mixed, and if, moreover, our current state of knowledge does not allow us to make them pure by adding in further factors, then it might seem that, on any account of causation, we cannot trust those correlations as indicators of causal conclusions.

But we should not dismiss our original correlations too quickly. At this point we need to make a further distinction. Suppose that both P(C/S) and P(C/-S) are mixed with respect to some Z: Z makes a further difference to the probability of cancer, among both smokers and nonsmokers. We need to distinguish between cases where Z *screens C off from S*, and those where it doesn't.

Z is a "screener-off" in this sense if P(C/S&Z) = P(C/Z) and P(C/S&-Z) = P(C/-Z). Once we divide the reference class into Z and -Z, it turns out that smoking makes no real difference to the probability of cancer after all. Intuitively, screening off happens when Z is a common cause of S and C. Suppose there is a gene which, on the one hand, makes people likely to smoke, and, on the other hand, predisposes them to cancer. Among people with the gene, the smokers are no more likely to get cancer than the nonsmokers; and similarly among those without the gene. The original smoking-cancer correlation turns out to be due entirely to the fact that smoking is itself a symptom of the presence of the gene, in that the gene is present more often among smokers than among nonsmokers.

Note that not every Z which renders an S-C correlation mixed is one which

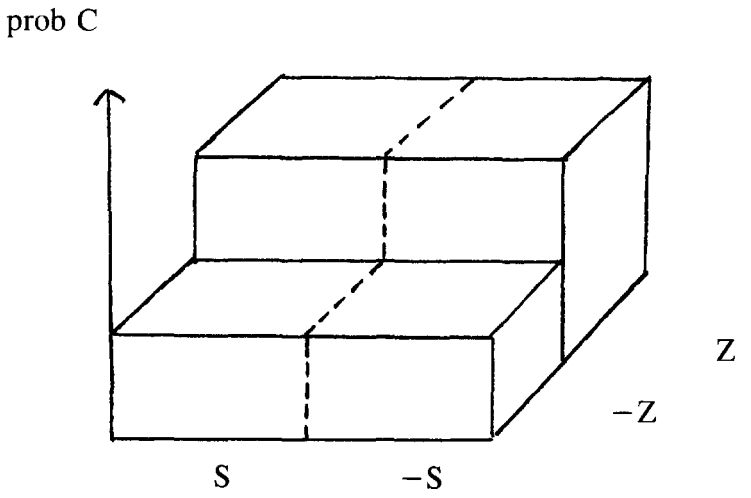

Figure 3. Z screens off C from S.

screens C off from S. A "screener-off" isn't just any old further factor that alters the probability of C. It alters the probability in a quite specific way: in changing it from P(C/S) to P(C/S&Z) it makes it equal to P(C/Z). A "mixer" just alters the probability somehow. A "screener-off" makes it equal to a specific number.

The fact that the existence of mixers does not imply the existence of screeners-off is important. A screener-off shows that S is in fact irrelevant to the probability of C. If Z screens C off from S, then partitioning by Z and -Z shows that the original division into S and -S is no longer necessary. But this irrelevance doesn't necessarily follow whenever Z renders the S-C correlation mixed. That S by itself leaves us with an inhomogeneous partition for C isn't yet any reason for thinking S isn't *part* of what is required for homogeneity. Even if there are unknown physical differences, Z, between people, which mean that some are more likely to get cancer than others, this doesn't mean that smoking doesn't make an extra difference as well. It may still be that P(C/S&Z) is greater than P(C/Z-by-itself). So an S-C correlation can be mixed and S can still be genuinely relevant to C: S can still be needed for the homogeneous partition which gives the pure probabilities of C.

Indeed we can say something stronger. Not only does mixedness allow that S *may* still be needed for the homogeneous partition, but, more strongly, if there aren't in fact any screeners-off, then an original S-C correlation shows that S *must* be part of the homogeneous partition for C. If there are no further factors which yield a partition which has no use for S, then S will still be needed when we get down to the final partition which gives us the pure probabilities of C.

Recall once more the probabilities with which I began. Obviously, as I have

already argued, these probabilities are mixed, both by Z's we can identify and, no doubt, by some we can't. But this doesn't mean that there is a Z which screens cancer off from smoking, or the grape harvest off from the August sun, or getting "6" 's off from the die being loaded. And I take it that in fact there aren't any such screeners-off. I accept that our original probabilities are mixed. But I don't think they are *spurious*. I don't think that there are any Z's which will show that the apparent causes are merely symptoms of the apparent effects, which will show that at bottom smoking (or sun, or being loaded) makes no real difference to cancer (or the harvest, or getting a "6").

5. Spuriousness and Statistical Research

How can we have any epistemological confidence that probabilities like the above aren't spurious, once we admit that they are mixed by unknown factors? Doesn't the admission that there are further unknown Z's which alter the probabilities always leave us with the danger that some of those Z's will be screeners-off? It seems as if we will only be able to dispel the epistemological danger of spuriousness if we can identify all the Z's and bring them explicitly into our analysis.

But note that to deal with the epistemological danger of spuriousness we need only take into account those other Z's *that S is itself associated with*, not all the other mixers. Only factors that are themselves associated with S are potential screeners-off. If S makes no difference when you do take Z into account, then it can only seem to make a difference when you don't take Z into account, if S itself happens with Z more often than without.

So to rule out spuriousness, we don't need to take all other Z's into account. We only need consider those that might be associated with S. I take it that this requirement is satisfied in our examples. We haven't identified all the factors that make a difference to the chance of cancer. But this doesn't mean that we haven't looked at all those that might themselves be associated with smoking. Thus, for instance, I presume that tobacco-sponsored researchers have checked to see if anxiety, which is itself associated with smoking, screens off smoking from cancer, and I presume that they have discovered it doesn't. But I expect that even the cigarette companies haven't bothered to take explicit account of, say, pre-natal developmental abnormalities of the lungs, for although these may well increase the chance of cancer, it is hard to see how they could possibly be associated with smoking.

The point I am making here is of fundamental importance for much medical, agricultural, biological, psychological, and sociological research. (Incidentally, it was my interest in the logic of this kind of research that got me interested in the connection between probabilities and causes in the first place.) Statistical data in such areas almost inevitably yield mixed probabilities, not pure ones. If mixing couldn't be separated from screening off without taking all relevant factors into

account, then causal conclusions in these areas would be impossible. So it is fortunate that the medical, etc., researchers don't need to take all factors into account, but are entitled to draw causal conclusions once they have explicitly considered all possibly "confounding" factors—that is, once they have explicitly considered all those factors which might themselves be probabilistically associated with the putative cause.

This of course is why "randomized" experiments are so important. If you can split your sample into cases which get the putative cause, and cases which don't, at random, then you can be sure that the other factors relevant to the effect will themselves be probabilistically independent of the putative cause. And this then means that you can be sure that any association between the putative cause and effect won't be spurious. (Of course you might get *sample* associations which were misleading about population probabilities. But that's different. Here and throughout I'm prescinding from evidential issues and assuming we're dealing with real probabilities. I intend spurious probabilities, and, for that matter, mixed probabilities, to be understood as real population probabilities. Spuriousness isn't a matter of misleading samples, but of misleading causal significance.)

We can't always conduct randomized experiments. Often there will be moral or practical reasons why we can't manipulate the data. In such cases we will need to conduct surveys instead, and to try to ensure that all possible confounding factors are taken into explicit account. This demand obviously raises difficulties that are avoided by randomizing. But there is no reason to suppose that this demand can't often be satisfied. Even if we can't experiment on smokers, we can still survey for all the factors that might be confounding the association between smoking and cancer, and, as pointed out above, can achieve some assurance that all potential confounders have been catered for. In principle, no doubt, there always remains a possibility that our causal conclusions are in error. But that is a feature of empirical research in general, not a danger that's peculiar to inferences from surveys.

6. The Importance of the Single Case

I said that my initial intention in this paper was to show how the deterministic view of causation can withstand the intuitive plausibility of the S-R view. So far I have distinguished between spurious and mixed cases of statistical relevance, and pointed out that spurious probabilities are misleading as to causes. But doesn't this simply add weight to the S-R view? After all, no S-R theorist has ever wanted to say that *any* case of probabilistic relevance establishes causation. (Even first-year sociology students don't want to say that correlation always equals causation.) According to Suppes, an initial correlation between S and C shows only that S is a *prima facie* cause of C. For S to be a *genuine* cause of C, there needs in addition to be no Z which screens C off from S.

But isn't that just what I have said? I have urged that we take smoking to cause cancer because, even though we believe the original correlation to be mixed, we don't think it's spurious. This looks just like Suppes. Smoking is a genuine cause, as well as a *prima facie* one, because nothing screens C off from S.

But in fact Suppes's view is untenable. Suppose we have this kind of set-up. A genetic factor, Y, increases the chance of cancer in all those people who possess it. Smoking also, and independently, increases the chance of cancer, both in people with Y and in those without Y. But it only does so in conjunction with some further unknown metabolic condition X (smoking makes no difference at all to your chance of cancer if you lack X). So you have an increased chance of cancer if you either have the gene Y, or smoke and have the metabolic factor X, or both. Suppose nothing else is probabilistically relevant to C. Then, while P(C/S) is mixed, by X and by Y, neither of these further factors screens C off from S (smoking makes an extra difference both among X's and among Y's). So, according to Suppes, S is a genuine cause of C. But now consider this case. Mr. Jones smokes, and gets cancer. Presumably, according to Suppes, his smoking then caused his cancer. But suppose Jones didn't have X, but did have Y. In that case it seems clear that, even though we might misjudge the situation, his smoking wasn't in fact any part of what caused his cancer.

This case ("the case of the misleading cigarettes") shows conclusively that we can't rest with the canonical statistical-relevance view. Even if there is a sense in which unscreened-off correlations are important for causation—as indeed they are—it's not true that A causes B whenever A and B are nonspuriously associated.

But while the case of the misleading cigarettes causes difficulties for Suppes's theory, it is not hard to see what we ought to say about such cases. Jones's example shows that causal relationships in particular cases depend on single case chances, not population probabilities. The chance of *Jones* getting cancer given that he smoked was no greater than the chance he would have had if he hadn't smoked, for he lacked the relevant metabolic factor X. In the population at large nothing screens smoking off from cancer, and so the association between smoking and cancer isn't spurious. But that doesn't matter to the analysis of Jones, given that *his* smoking was irrelevant to *his* chance of cancer.

The moral of the story is that causal facts about particular cases depend on *pure* probabilities. It's not enough to make sure that the probabilities we are dealing with are not spurious. For as long as they remain mixed, we are in danger of being misled about the particular case. We need, so to speak, to get down to a homogeneous partition of the reference class. For unless we know which homogeneous cell Jones himself is in, we are in danger of attributing his chance of cancer to factors which matter in other cells, but not in his.

I might now seem to be taking back the points made in sections 3–5. There I said that we didn't need purity for causal conclusions: mixedness was quite all right, as long as it wasn't spurious. But now I'm saying that we do need purity.

The solution to this conundrum, such as it is, is that there are two kinds of causation at issue. In this section I have been focusing on conclusions about specific cases: did *Jones's* smoking cause *his* cancer? But in earlier sections it was a general, population-relative question: is smoking *a* cause of cancer? It was this latter, population-relative question which I argued earlier can be decided by mixed nonspurious probabilities. For the population-relative question isn't about any particular cell of the homogeneous partition, but about the overall structure of that partition: in effect it is the question of whether smoking is relevant to *any* cell in the homogeneous partition. The point made earlier was that a nonspurious mixed correlation is necessary and sufficient for a positive answer to this question.

So let us distinguish two kinds of causation: "single case causation" and "population causation." The population notion is a perfectly natural notion of causation. It is the notion that we have in mind when we conclude from the statistics that smoking causes cancer. And, as we shall see in section 9, it is, give or take a bit, the notion that is relevant to rational decisions about action. But, even so, single case causation is clearly the more basic notion. In cases like the misleading cigarettes, our intuitions are unequivocably that the smoking didn't *cause* the cancer, even though smoking is a perfectly good population cause of cancer. And indeed population causation is definable in terms of single case causation: to say that smoking is *a* cause of cancer is simply to say that smoking is *sometimes* a single case cause of cancer (namely, when conjoined with the relevant background factors).[2]

7. The Compatibility of Probabilistic Intuitions with a Deterministic View of Causation

In the last section I showed that Suppes's original formulation of the S-R view won't do as an account of the basic single case notion of causation. But this scarcely eliminates the S-R view as such. For the obvious response is to tighten up the S-R view in light of the distinction between single case causation and population causation. Thus, for instance, the S-R view could be put forward specifically as an analysis of the primary notion of single case causation. The S-R view would then be the theory that an increase in probability amounts to a genuine single case cause, as opposed to a prima facie single case cause, as long as that increase in probability is an increase in the pure chance of the result.

On this account, then, the kind of mixed but nonspurious probabilities discussed in sections 3–5 would cease to be of any direct relevance to the basic notion of single case causation. Such mixed nonspurious probabilities would relate only to the derivative notion of population causation. Single case causation would itself be explained entirely in terms of pure probabilities.

From now on I shall understand the S-R view as this thesis about single case causation. The switch to the single case is the obvious way for an S-R theorist to respond to the case of the misleading cigarettes. But now another problem faces the S-R theorist. Once it is allowed that mixed nonspurious probabilities are of no special significance to causation as the S-R theory conceives it, then the initial intuition which seemed to support the S-R view ("Doesn't the smoking-cancer correlation indicate that smoking *causes* cancer?") ceases to count specifically in its favor. For it now turns out that this initial intuition is entirely compatible with a deterministic view of causation.

The crucial point to note here is that the analysis of the last section, about the difference between population and single case causation and about the real significance of mixed nonspurious probabilities, is all *perfectly consistent with the assumption that everything is determined.* To see this, imagine that prior circumstances always leave any individual with either a unitary or a zero chance of cancer. More specifically, imagine that the gene (Y), or smoking in conjunction with the metabolic factor (S&X), or both, make it certain that you will get cancer, and that lacking both Y and S&X makes it certain you won't get cancer. That is, imagine that: S&X or Y $\leftrightarrow$ C.

Now in this situation the probability of cancer given smoking will still in general be less than one.[3] Unless Y or X is somehow ensured by smoking, not all smokers will get cancer. What is more, the probability of cancer with smoking will be greater than the probability of cancer without smoking, and both these probabilities will be mixed, and nothing will screen off the difference between them. And all this will show as before that smoking is a population cause of cancer. (Though again we should be careful not to fall into the trap of the misleading cigarettes: that is, we shouldn't take it for granted that any particular smoker's cancer was caused by his or her smoking.)

So everything remains just as it was in the last section. It is easy to see why. To assume determinism is simply to assume that there are only two kinds of cell in the homogeneous partition: those in which cancer's probability is one, and those in which it is zero. So determinism is simply a special case of the situation considered so far, and all the arguments remain valid.

Indeed under the assumption of determinism the nature of the inference from a nonspurious mixed correlation to a population causal conclusion is particularly perspicuous. Let me spell the point out. Among the nonsmokers we'll get cancer if and only if Y. Among the smokers, on the other hand, we'll get cancer if and only if X *or* Y. Which means that cancer will be more likely among smokers than nonsmokers if and only if P(X or Y) is greater than P(Y)—that is, if and only if the background factors which determine cancer together with smoking aren't an empty set.

Actually, that's a bit quick. Even if we leave measure-theoretic and modal niceties to one side (see my 1985a, 73), there's the obvious point that cancer can

get to be more likely among smokers than nonsmokers, not because of any nonempty X, but because some elements in Y are disproportionately common among smokers (imagine that Y includes the carcinogenic gene which also induces people to smoke). But this is easily remedied. Consider separately classes of people alike in any such elements of Y, and see whether smoking still makes cancer more likely within them. (Or, to slip back into my original terminology, we need to make sure that the association between smoking and cancer isn't spurious, by taking into explicit account relevant factors themselves associated with smoking.)

The point of all this has been to show that, even if we were dyed-in-the-wool determinists who thought that everything was made certain or impossible by prior circumstances, we would still be very interested in population probabilities that differed from zero and one, and would still be inclined, quite rightly, to infer population causal conclusions from mixed but nonspurious probabilities.

And this now shows how the intuition that increased probabilities establish causation can be accounted for in a way that gives no special support to the theory that single case causation itself is probabilistic. For even if you believed in determinism, and therefore took it for granted that all single case causation was deterministic, you would still be served well by the intuition that an increased probability indicates population causation.

What is more, it seems highly likely that determinism is in fact the basis of that intuition, historically speaking. After all, the view that the world is at bottom indeterministic has only recently been accepted by informed common sense. (Most people probably still believe that macro-events like cancer, or crop yields, or coin tosses, are always determined when they occur.) No doubt writers like Suppes and Salmon were inspired to formulate the S-R view by the indeterminism of modern physics. But the reason their claims struck such an intuitive chord in their readers is surely quite different, namely, our implicit grasp of the fact that in a deterministic universe S can only be nonspuriously correlated with C if S figures among a conjunction of factors that determines C.

This isn't of course an argument against the S-R view. But it does show that the S-R view can't simply appeal to the intuitive connection between probability increases and causes. For this intuitive connection draws most of its strength, not from the S-R view, but from the contradictory view that causation is deterministic.

8. The Deterministic Causation of Chances

Perhaps I can give a genetic account of the power of the intuition that statistical relevance implies causation, by appealing to the long-standing presupposition that macro-events are always determined. But in fact this presupposition is false.

Quantum mechanics tells us that the readings on geiger counters and like devices are undetermined macro-events. And there is no reason not to allow that even such nonlaboratory macro-events, like getting cancer or tossing a coin, might also be undetermined: perhaps they depend on the quantum mechanical breaking of bonds in the DNA molecules in lung tissue cells, or, again, in the breaking of bonds in the air molecules that the dice collide with.

If determinism were indeed true, then nonunitary probabilities would always be due to our ignorance of all relevant factors, and it would be sensible to insist that such probabilities were merely indirect evidence for underlying deterministic single case causes. But this scarcely decides what we should say given that determinism *isn't* true, and that nonunitary probabilities can perfectly well be pure reflections of complete information.

Let me finally face the issue. Suppose cancer is an indeterministic phenomenon. And suppose we are not misled about Jones. Jones does have X. So his smoking does increase the chance of his getting cancer. And then he gets cancer. Surely I have to say that his smoking caused his cancer.

But, as I indicated in the Introduction, there is another possible line here. This allows that something causal is going on. But it denies that Jone's *cancer* is caused. Rather what is caused is his increased *chance* of cancer. And this then is consistent with the view that all causation is deterministic: for, by hypothesis, whenever smoking (and X) occurs, then that increased chance of cancer is determined.

This line might seem unnatural. Doesn't intuition tell us that in such cases Jones gets *cancer* itself because of his smoking? What's the point of insisting that the only real effect is his increased *chance* of cancer, apart from a misplaced hankering for the old deterministic metaphysics?

But I see no reason to accept that we have any real intuitions about situations like this. We are supposing, for the sake of the argument, that we have complete knowledge about Jones, that we aren't being misled by his smoking. But in the real world we are unfamiliar with this kind of situation. Nearly all the actual probabilities different from zero and one that we use in our everyday lives are indubitably mixed, as I stressed in sections 3–5. We recognize intuitively that such mixed probabilities are a sound guide to population causal conclusions (as long as they aren't spurious). But this intuitive recognition, that our incomplete knowledge entitles us to some kind of population causal conclusion, is quite consistent with either analysis of single case causation: either that single case causation is a matter of prior circumstances determining later chances, or that single case causation is a matter of the indeterministic causation of actual results. And so, it seems to me, our intuitive responses in situations of incomplete knowledge leave it quite open which alternative we would adopt if we ever were in the unfamiliar situation of complete knowledge.

9. Rational Action

So far I have argued that our existing intuitions are consistent with the view that single case causation is deterministic. But ought the matter to rest on existing intuitions? Perhaps there are some positive *arguments* for admitting indeterministic single case causation.

In this section I want to consider one such possible argument for indeterministic single case causation. (Let me take the "single case" as read from now on, unless I say otherwise.) This argument derives from the connection between causation and rational action, and goes like this: it is rational to do A in pursuit of B just in case you believe A causes B; but normative decision theory tells us that it is rational to do A in pursuit of B just in case you believe Prob (B/A) > Prob (B/−A); so doesn't it follow that A causes B just in case Prob (B/A) > Prob (B/−A)? (See Mellor forthcoming.)

There is more to this argument than meets the eye. Let us start by looking more closely at the second premise, that it is rational to do A in pursuit of B just in case you believe Prob (B/A) > Prob (B/−A). As is now well known, this premise needs qualification. We know that the probability of avoiding driving accidents (B) if you are a house-owner (A) is greater than the probability of avoiding them if you are not: Prob (B/A) > Prob (B/−A). But this doesn't mean that it is rational to buy a house in order to avoid accidents.

In this case there is no doubt some underlying character type, C, say, which both influences people to buy houses and leads them to drive carefully. Either you have this character type or you don't, and in neither case is buying a house going to make any extra difference to your chance of avoiding an accident. Which is why it isn't sensible to buy a house in order to avoid accidents.

The original probabilities here, Prob (B/A) and Prob (B/−A), are mixed, in that the further factor C is also relevant to the probability of B. Moreover, C is not just a mixer, but is also a screener-off: Prob (B/A&C) = Prob (B/−A&C), Prob (B/A&−C) = Prob (B/−A&−C). What does all this have to do with the rationality of action? How should we qualify decision theory to cope with this case?

One possible response here would be to insist that you should only act on probabilities when you believe them to be pure. The reasoning would be that, as long as your probabilities are mixed, you are in danger of being misled about your actual situation. You can know that more smokers get cancer than nonsmokers. But if this is a mixed probability, you might be a not-X: you might be the kind of person in whom smoking makes no difference to the chance of cancer. In which case there wouldn't be any point in your giving up smoking to avoid cancer.

But the requirement that you believe your probabilities to be pure is surely too strong a condition on rational action. There is a sense in which you might always be misled into wasting your time if your probabilities are mixed. But that doesn't

mean that your action would be *irrational*. After all, to repeat the point yet again, nearly all the probabilities we come across in everyday life are indubitably mixed, by factors that we don't know how to identify. The smoking-cancer correlation is just such a probability. So clearly asking for pure probabilities is too strong. If it were irrational to act on probabilities you believed to be mixed, nobody would ever have been right to give up smoking to avoid cancer.

It's not mixed probabilities that are a bad basis for action, but spurious ones. Think of it in terms of homogeneous partitions of reference classes. If your probability is mixed by factors you can't identify, then you don't know which cell of the partition you are in (you don't know whether you have X or not), and so you don't know what difference your action will actually make to the chance of the desired outcome. But, still, you *may* be in a cell where your action makes a difference, and this in itself gives you reason to act. But if your probability is spurious, then your action *can't* make a difference, for *whichever* cell you are in, your action will be rendered irrelevant to the desired outcome by the screener-off (either you have C or not, and either way your house buying won't make any further difference to your accident-proneness).

10. Quantitative Decisions

So the moral is that it is perfectly rational to act on probabilities that you recognize to be mixed, as long as you don't think they are spurious as well. Can we be more specific? So far my comments on decision theory have been entirely qualitative. But normative decision theory deals with numbers. It tells you how *much* probabilistic beliefs should move you to act. You should act so as to *maximize* expected utility. The desirability of an action should be proportional to the *extent* to which it is believed to make desired outcomes likely.

Can't we just say that probabilities will be quantitatively suitable for expected utility calculations as long as you believe they aren't spurious? The thought would be this. As long as you believe your probabilities aren't spurious, the differences between Prob (B/A) and Prob (B/-A) can be thought of as a weighted average of the difference A makes to B across all the different cells of the homogeneous partition. You don't know which cell you are actually in. You might be in a cell where A makes no difference. You might even be in a cell where A makes B less likely. But, even so, the overall difference between Prob (B/A) > Prob (B/−A) tells you how much difference A makes on weighted average over all the cells you might be in.

But this won't do. So far I have understood spuriousness as an entirely on-off matter. Spuriousness has been a matter of complete screening off, in the sense of a correlation between putative cause A and putative effect B disappearing entirely when we control by some further X. But spuriousness also comes in degrees. A confounding background factor can distort a correlation, without its

being the case that the correlation will be completely screened off when we take that factor into account. Rational action needs to be sensitive to the possibility of such partial spuriousness.

Let me illustrate. Suppose once more that there is a gene which conduces, independently, to both smoking and cancer; but now suppose also that smoking makes a slight extra difference to the chance of cancer: both among those with the gene, and among those without, the smokers are slightly more likely to get cancer. In this case the gene won't entirely screen smoking off from cancer. Controlling for the gene won't reduce the correlation between smoking and cancer to zero. Yet the extent to which smoking is associated with cancer in the overall population will be misleading as to its real influence, and therefore a bad basis for decisions as to whether to smoke or not. Smoking will at first sight seem to be much more important than it is, because of its positive association with the more major cause of cancer, possession of the gene.

Technically we can understand the situation as follows. Prob (B/A) and Prob (B/−A) are indeed weighted averages. But they are weighted by the inappropriate quantities for expected utility calculations. Let us simplify by supposing that X is the only other factor apart from A relevant to B. Now,

(1) Prob (B/A) = Prob (X/A) Prob (B/A&X) + Prob (−X/A) Prob (B/A&−X),

and

(2) Prob (B/−A) = Prob (X/−A) Prob (B/X&−A) + Prob (−X/−A) Prob (B/−X&−A)

This is the sense in which Prob (B/A) and Prob (B/−A) are indeed weighted averages of the probability that A (respectively, not−A) gives B in the "X−cell," and the probability that A (not−A) gives B in the "not−X" cell. But the weighting factors here, Prob (X/A) and Prob (−X/A) (respectively, Prob (X/−A) and Prob (−X/−A)), aren't what we want for rational decisions. They depend on the extent to which A is associated with X, and so mean that the difference between Prob (B/A) and Prob (B/-A) reflects not just the influence of A on B, but also the correlation of A with any other influence on B. In the extreme case, of course, this can make for an overall difference between Prob (B/A) and Prob (B/−A) even though A makes no real difference at all: even though Prob (B/A&X) = Prob (B/A&−X), and Prob (B/A&−X) = Prob (B/−A&−X), and X entirely screens off A from B. But the present point is that, even without such complete screening off, any association between A and X will confound the correlation between A and B and make it seem as if A has more influence than it does.

What does this mean in practical contexts? Are quantitative utility calculations only going to be sensible when we have complete knowledge and pure probabili-

ties? Not necessarily. For note that there is nothing wrong with the weighted average argument if we use the right weights, namely P(X) and P(−X), and so really do get the weighted average of the difference A makes in the X-cell and the not-X-cell respectively. That is, the right quantity for utility calculations is

(3) Prob (X) [Prob (B/A&X) − Prob (B/−A&X)] + Prob (−X) [Prob (B/A&−X) − Prob (B/−A&−X)]

In the special case where A is not associated with X, the weighting factors in the earlier equations (1) and (2) reduce to P(X) and P(−X), and the difference between P(B/A) and P(B/−A) therefore reduces to the requisite sum (3). But if there is an association between A and X, then we have to "correct" for this confounding influence by replacing the conditional weighting factors in (1) and (2) by the correct P(X) and P(−X).

To illustrate with the smoking-cancer-gene example, you don't want to weight the difference that smoking makes within the "gene-cell" by the respective probabilities of *smokers* and *nonsmokers* having the gene, as in (1) and (2), because that will "bump up" the apparent influence of smoking on cancer in line with the positive likelihood of smokers having been led to smoke by the gene. The issue, from the agent's point of view, is precisely whether or not to smoke. And so the appropriate quantity for the agent is the probability of *anybody* having the gene in the first place, whether or not they smoke, not the probabilities displayed by smokers and nonsmokers.

The practical upshot is that anybody interested in quantitative utility calculations needs to take into explicit account any further influences on the result that they believe the cause (action) under consideration is associated with. If you don't believe there are any possible confounding influences, then you can go ahead and act on Prob (B/A) − Prob (B/−A). But if you do think there are associations between other causes X and A, then you will need to turn to the "corrected" figure (3).

11. Causal and Evidential Decision Theory

In the last two sections I have been considering how rational decision theory should respond to the danger of spuriousness. This topic has been the subject of much recent debate. The debate was originally stimulated by Newcomb's paradox (see Nozick 1969), which is in a sense an extreme case of spuriousness. But it has become clear that the underlying problem arises with perfectly straightforward examples, like those I have been discussing in the last two sections.

Philosophers have fallen into two camps in response to such examples: evidential decision theorists and causal decision theorists. Causal decision theorists argue that our decisions need to be informed by beliefs about the causal structure

of the world (see Lewis, 1981, for a survey of such theories). Evidential decision theorists, on the other hand, try to show that we can manage with probabilistic beliefs alone: they feel that we ought not to build philosophically dubious metaphysical notions like causation into our theory of rational decisions if we can help it (see Eells 1982; Jeffrey 1983).

At first sight it might seem that I am on the side of evidential decision theory. All of my analysis in the last two sections was in terms of various conditional probabilities, as in equations (1)–(3) of the last section. But this is misleading. For my recommended decisions require an agent to take a view about *spuriousness*, and spuriousness, as I have defined it, depends on an underlying metaphysical picture. (For any sort of effect E, there is a set of factors which yield an *objectively* homogeneous partition of the reference class with respect to E; spuriousness then depends on whether any of *those* factors screen C off from E).

Given the general tenor of evidential decision theory, and in particular given the structure of the "tickle defense" (to be discussed in a moment), it is clear that evidential decision theorists would find my appeal to the notion of objective probability as objectionable as the appeal to the notion of causation. From their point of view my approach would be just as bad as causal decision theory – I'm simply using the notion of objective probability to do the work of the notion of causation.

I am inclined to see things differently. I would say that the possibility of substituting objective probabilities for causes makes causes respectable, not objective probabilities disreputable. And in section 13 I shall begin exploring the possibility of such a reduction of causation to probability at length. But first let me go into a bit more detail about the different kinds of decision theory.

The underlying idea behind evidential decision theory is that we can manage entirely with *subjective probabilities*, that is, with our subjective estimates of how likely one thing makes another, as evidenced in our betting dispositions. This commitment to subjective probabilities is then combined with a kind of principle of total evidence: we should conditionalize on everything we know about ourselves (K), and we should then perform act C in pursuit of E according as $P(E/C.K) > P(E/K)$.

But evidential decision theory then faces the difficulty that the above inequality may hold, and yet an agent may still believe that the correlation between C and E within K is (to speak tendentiously) objectively spurious. And then of course it doesn't seem at all rational to do C in pursuit of E. If I think that some unknown but objectively relevant character trait screens house-buying off from lack of car accidents, then it's obviously irrational for me to buy a house in order to avoid car accidents.

The standard maneuver for evidential decision theorists at this point is some version of the "tickle defense" (see Eells, ch. 7). In effect defenders of evidential decision theory argue that an agent's total knowledge will always provide a reference class in which the agent believes that the C-E correlation is *not* spurious.

The underlying reasoning seems to be this: (a) spurious correlations always come from common causes; (b) any common cause of an action type C and an outcome E will need, on the "C-side," to proceed via the characteristic reasons (R) for which agents do C; (c) agents can always introspectively tell (by the "tickle" of their inclination to act) whether they have R or not; and so (d) they can conditionalize on R or -R), thereby screening C off from E if the correlation is indeed spurious, and so avoid acting irrationally. To illustrate, if the house-buying/car-safety correlation is really due to causation by a common character trait, then I should be able to tell, by introspecting my house-buying inclinations, whether I've got the trait or not. And so the probabilities I ought to be considering are not whether house-buyers as such are more likely to avoid accidents than nonhouse-buyers, but whether among people with the character trait (or among those without) house-buyers are less likely to have accidents (which presumably they aren't).

This is all rather odd. The most common objection to the tickle defense is that we can't always introspect our reasons. But that's a relatively finicky complaint. For surely the whole program is quite ill-motivated. The original rationale for evidential decision theory is to avoid metaphysically dubious notions like causation or objective probability. But, as I hope the above characterization makes clear (note particularly steps (a) and (b)), the tickle defense only looks as if it has a chance of working because of fairly strong assumptions about causation and which partitions give objectively nonspurious correlations. It scarcely makes much sense to show that *agents* can always manage without notions of causation and objective probability, if our *philosophical* argument for this conclusion itself depends on such notions.

Perhaps the defenders of evidential decision theory will say they are only arguing *ad hominem. They* don't believe in objective spuriousness, common causes, etc. It's just that their opponents clearly have such notions in mind when constructing putative counter-examples like the house-buying/car-safety story. And so, the defenders of the evidential theory can say, they are merely blocking the counter-examples by showing that even assuming their opponents' (misguided) ways of thinking of such situations, there will still always be an evidentially acceptable way of reaching the right answer.

But this now commits the evidential decision theorist to an absurdly contorted stance. If evidential theorists really don't believe in such notions as causation, objective spuriousness, etc., then they are committed to saying that the mistake you would be making if you bought a house to avoid car accidents would be (a) that you hadn't introspected enough and therefore (b) that you hadn't conditionalized your house-buying/car-safety correlations on characteristics you could have known yourself to have. But that's surely a very odd way of seeing things. You don't need to be introspective to avoid such mistakes. You just need to avoid acting on patently spurious correlations. Pre-theoretically, it's surely their insensi-

tivity to manifest spuriousness that makes us think that such agents would be irrational, not their lack of self-awareness. It seems to me that there must be something wrong with a theory that denies itself the resources to state this simple fact.

One can sympathize with the original motivation for evidential decision theory. The notion of causation is certainly philosophically problematic. And perhaps that does give us some reason for wanting the rationality of action not to depend on beliefs about causal relationships. But, now, given the way I have dealt with rational action, agents don't need causal beliefs, so much as beliefs about whether certain correlations are objectively spurious or not. The fact that evidential decision theorists feel themselves driven to the "tickle defense" shows that they wouldn't be happy with the notion of objective spuriousness either. But putting the alternative in terms of objective probabilities now places evidential decision theory in a far less sympathetic light. For even if the notion of objective probability raises its own philosophical difficulties, modern physics means that we must somehow find space for this notion in our view of the world, and so removes the motivation for wanting to avoid it in an account of rational action. Moreover, if the cost of keeping objective probabilities out of rational decision theory is the contortions of the "tickle defense," then we have a strong positive reason for bringing them in.

I now want to leave the subject of evidential decision theory. The only reason I have spent so long on it is to make it clear that, despite initial appearances, the approach I have adopted is quite different, and indeed has far more affinity with causal decision theory. Let me now consider this latter affinity. On my account rational action requires you to believe that, even if your correlations are mixed, they are not spurious. If you believe your correlations are spurious, to any degree, then you need to correct them, in the way indicated in the previous section: you need to imagine the reference class partitioned into cells within which such spuriousness disappears, and then to average the "within-cells" correlations, weighted by the probability of your being in each cell.

According to causal decision theory, it is rational to act if you believe that your correlations reflect a causal, and not merely an evidential, connection between your action and the desired result. If you believe the correlations are evidential, then you need to consider separately all the different hypotheses about the causal structure of the world you believe possible, and then average the difference that the action makes to the chance of the result under each hypothesis, weighted by the probability that you attach to each hypothesis.

I don't think there is any real difference here. I think that the two approaches simply state the same requirement in different words. This is because I think that facts of causal dependence can be entirely reduced to facts about probabilities in objectively homogeneous partitions of reference classes. But this is itself a contentious thesis. There are various difficulties in the way of this reduction, many

of which I have been slurring over so far. Most of the rest of the paper will be devoted to dealing with them.

Note that this issue of reduction is independent of the debate between the S-R and traditional deterministic views of causation. The idea I want to explore (I shall call it the "reductionist thesis" from now on) is that we have causal dependence of E on C if and only if C and all the other probabilistically relevant factors present define a homogeneous cell of the reference class which yields a higher probability for E than is yielded those other relevant factors alone; or, again, if and only if the chance of E given C and all the other relevant factors present is higher than the chance E would have had given those other factors but without C.[4] But now suppose that this reductionist thesis were granted. This would still leave it quite open whether in such cases we should say that C (indeterministically) caused *E*, or whether we should say that C (deterministically) caused the increased *chance* of E.

I'm not going to have much more to say about this latter issue. It seems to me that by now these are pretty much just two different ways of talking (and in discussing the reductionist thesis I shall adopt both indiscriminately). But an earlier argument for the S-R view has been left hanging in the air. Let me briefly deal with this before turning to the general issue of reduction.

12. Action and Causation Again

The argument in question is the one from the beginning of section 9: (a) it is rational to do A in pursuit of B just in case you believe A causes B; (b) it is rational to do A in pursuit of B just in case you believe $P(B/A) > P(B/-A)$; so (c) A causes B just in case $P(B/A) > P(B/-A)$.

I have shown that the second premise (b) won't do as it stands. Not all correlations are a good basis for action. It doesn't matter in itself if a correlation is believed to be mixed. But a correlation is disqualified as a basis for action if it is believed to be spurious. Before we act we need to take into account all the factors that we believe to be confounding the association between A and B, and adjust the correlation accordingly.

It might seem as if this now means that I can respond to the argument at hand as I originally responded (in sections 7 and 8 above) to the initial intuition favoring the S-R view. That is, can't I point out that all the probabilities we actually act on are undoubtedly mixed? We recognize that such probabilities had better not be spurious. But we also recognize that they don't need to be pure. So our intuitions about when it is and isn't rational to act are quite consistent with the supposition that all events are determined and that the only reason we have probabilities other than nought and one is that we are ignorant of various relevant (but non-confounding) factors. And not only are our intuitions so consistent with determinism, they are no doubt inspired by it, since until recently determinism was

built into informed common sense. So we can scarcely appeal to such intuitions to decide against a deterministic view of causation.

But this argument won't serve in the present context. For the S-R theorist isn't now appealing to mere intuitions about causation. Rather the appeal is to *facts*, so to speak, about when it's rational to act, over and above any intuitions we may have on the matter. This means that the S-R theorist can now insist that the relevant situation is one where an agent believes that a result is genuinely undetermined. Maybe we don't have any immediate causal intuitions about such indeterministic set-ups, since until recently we didn't believe there were any. But that doesn't stop there being a fact of the matter as to how one ought to act in such situations.

And here the S-R theorist is clearly on strong ground. For there is no question but that knowledge of objective nonunitary chances can be relevant to rational action. If I believed that the effect of not smoking, when all other relevant factors are taken into account, is to increase the chance of avoiding cancer, but without determining it, then obviously this would give me a reason to stop smoking.

So the fact that we were all determinists till recently is irrelevant to the argument from rational action. The issue is not *why* we think that it's rational to act if and only if (nonspuriously) $P(B/A) > P(B/-A)$. Rather the point is that it *is* so rational (and in particular that it is so rational even if P(B/A)'s being less than one isn't just due to our ignorance of the relevant determining factors).

But there is still room to resist the S-R view. Even if we concede premise (b), we can still question premise (a). Premise (a) says it is rational to do A in pursuit of B just in case you believe A causes B. But why not say instead that it is rational to do A in pursuit of B just in case you believe A causes an increased chance of B? This will enable us to accommodate all the relevant facts about rational action, while still preserving a deterministic view of causation.

This argument is clearly in danger of degenerating into triviality. But let me just make one observation before preceeding. It might seem *ad hoc* for the traditional theorist to start fiddling with premise (a) when faced by indeterminism. But note that the S-R theorist also has to do some fiddling with (a) in the face of indeterminism. The S-R theorist can't simply say that A causes B whenever it's rational to do A in pursuit of some B. For A can make B more likely, and yet B might not occur. The S-R notion of causation isn't just that A increase the chance of B, but that A increase the chance of B, *and* B occurs. So the S-R theorist has to formulate (a) in some such form as: A causes B just in case it's rational to do A in pursuit of B, *and* B occurs; or, again, it's rational to do A in pursuit of B just in case you're in the kind of situation where A *might* cause B. It's not clear to me that these formulations are any more satisfactory than the deterministic alternative suggested in the last paragraph, according to which it is rational to do A in pursuit of B just in case A invariably causes an increased chance of B.

13. The Metaphysics of Probability

In her (1979) Nancy Cartwright argues against the reducibility of causal relationships to laws of probabilistic association. Her argument depends on the point that probabilistic relationships only indicate causal relationships if they don't get screened off when we conditionalize on relevant background factors – causation demands nonspurious associations, not just any associations. However, the idea of nonspuriousness requires a specification of the class of background factors which need to be taken into account. Cartwright argues that this can only be given as the class of *causally* relevant factors.

So Cartwright allows that a causal relationship is a probabilistic association that doesn't get screened off by any causally relevant factors. This gives us a relationship between causal and probabilistic notions. But the appearance of the notion of "causal relevance" on the right-hand side of this relationship clearly rules it out as a reduction of causation.

Cartwright's argument is often endorsed in the literature (see Eells and Sober 1983, 38; Eells and Sober 1986, 230). But it seems to me that it is easily answered. Why not just say that the factors that need to be taken into account are all those which are *probabilistically*, rather than causally, relevant to the result? This would accommodate the possibility of spuriousness, but without rendering the proposed reduction circular.

What is a "probabilistically relevant" factor for some result E? It's any property which, in conjunction with certain other properties, is relevant to the chance of E. That is, it's any K such that there exists an L such that P(E/K.L) and P(E/-K.L) are pure and unequal.

Putting it like this makes it clear that we don't really need a restriction on the set of factors relevant to spuriousness in the first place. For conditionalizing on probabilistically *ir*relevant factors isn't going to show any probabilistic associations to be spurious, since by definition irrelevant factors don't make any difference to probabilities. So we may as well simply say that a probabilistic association indicates a causal relationship as long as there isn't any background factor which screens it off.

Will we ever have any causal relationships, if a causal relationship is disproved by *any* factor which screens the putative cause C off from the effect E? Surely there will always be *some* way of categorizing things that equalizes the proportions with which E is found with and without C. (See Cartwright 1979, 434.)

At first sight this objection might seem plausible. But it can be countered if we take care to make the distinction between the epistemology and the metaphysics of probability (as Cartwright herself notes, though she has her doubts about the distinction). Certainly if we are dealing with sample frequencies there will always be some way of dividing the sample into two parts that equalizes the relative frequency with E is found with and without C. But that's quite differ-

ent from the idea that there's always a property that will render C irrelevant to the *chance* of E.

I take it that there are real chances in the world: chances of certain properties being instantiated, in situations defined by certain other properties. Throughout this paper I have intended "probabilities" to be understood either as chances, or as chance mixtures of chances (that is, as the average of chances in different homogeneous cells weighted by the probability of being in each cell).

Probabilities (chances and chance mixtures of chances) manifest themselves in, and are evidenced by, relative frequencies in finite samples. The relationship between probabilities and such frequencies is a deep and difficult issue. But, however that issue is to be resolved, it is clear that not every relative frequency in every sample corresponds to a real probability. And, in particular, it is clear that it doesn't follow, just because sample correlations are always screenable off, that real correlations always are.

It might be objected that by helping myself to an ontology of chances and probabilistically relevant properties, I am begging all the interesting questions. Isn't having chances and the properties they involve tantamount to knowing about the causal structure of the world? What is the difference between probabilistically relevant properties and straightforwardly causally relevant ones?

In a sense I am sympathetic to this complaint. After all, I want to show that the causal facts reduce to probabilistic ones. But this doesn't mean that their relation is trivial, or that there's no point in trying to spell it out. If the reduction is possible, then there is a sense in which causal facts are built into probabilistic facts. But it's certainly not obvious at first sight that this is so.

14. Causal Chains

In this section and the next I want to look at some difficulties to do with the relationship between causation and time.

So far I've been cheating. I've in effect assumed that there are just two times, "earlier" and "later." The effect E happens "later." A number of factors are present "earlier." The reductionist thesis was that an earlier factor C is a single case cause of a later E just in case the chance of E was higher than it would have been if C had been absent and all other relevant earlier factors had been the same.

But of course there aren't just two times, but a whole continuum. And in any case it's not clear why causes should always happen earlier and effects later.

In this section I want to look at a difficulty which arises as soon as we admit that there are more than two times, and even if we continue to assume that causes must precede their effects. In the next section I shall say something about what happens if we admit the possibility of effects preceding their causes.

As soon as we allow that there can be relevant factors temporally intermediate between C and E, there is a difficulty about an earlier C *ever* being a cause of

a later E. To be a cause means that you have to make a difference to the chance of E when *all* other relevant factors are taken into account. But now suppose that D is, intuitively speaking, causally intermediate between C and E: C causes E *by* causing D. For example, smoking causes cancer by getting nicotine into the lungs. D is clearly a relevant factor, if C is. But now, according to the reductionist thesis as so far stated, D is going to stop C counting as a cause of E. For C won't make any further difference to the chance of E once we take D into account. Given that you've got nicotine in your lungs, the fact that you smoke doesn't make you *more* likely to get cancer. And similarly if you haven't got nicotine in your lungs (imagine you are a rigorous non-inhaler) smoking won't make you more likely to get cancer either. (I'm assuming here for simplicity that nicotine is the only route by which smoking causes cancer.) The presence or absence of nicotine screens the smoking off from the cancer. And so, according to the reductionist thesis, the nicotine seems to stop smoking from causing cancer. But the argument is quite general. We seem forced to the undesirable consequence that nothing ever causes anything via intermediate causes.

It won't do to say that we shouldn't control for factors temporally intermediate between C and E. For perhaps C isn't in fact a genuine cause of E, but only appears to be so because it is associated with (though not the cause of) some real later cause D. And then it is precisely that C doesn't make a difference when we conditionalize on D that should stop it counting as a genuine cause.

In Cartwright's eyes this provides an additional reason why we can't reduce causes to probabilities (Cartwright 1979, 424). Her original complaint was that we needed to specify the background factors to be taken into account as the set of "causally relevant" factors. I have answered that complaint by arguing that we may as well take *all* background factors into account. But now it seems that we need a further qualification. We shouldn't take all background factors into account after all, but only those which aren't causally intermediate between C and E, lest we end up ruling out all earlier C's as causes of later E's. But now this further qualification threatens to undermine the proposed reduction once more, since as before it seems that we need causal terminology (in particular, the notion of causal intermediacy) to explain which probabilistic relationships indicate causation.

I think there is a way out here. We need to distinguish between direct and indirect causes, and to define the latter in terms of the former.

Let us imagine that the times between C, at t_o, and E, at t_k, consist of a series of discrete instants, $t_1, t_2, \ldots, t_{k-2}, t_{k-1}$. (I shall relax the assumption of discreteness in a moment.)

Then we can say that a factor A at any T_1 is a *direct cause* of some B at the next instant, t_{1+1}, just in case the chance of B given A and all other factors present at t_1, or earlier, is greater than the chance of B given those other factors alone.

Then we can define a *causal chain* as a sequence of events at successive times

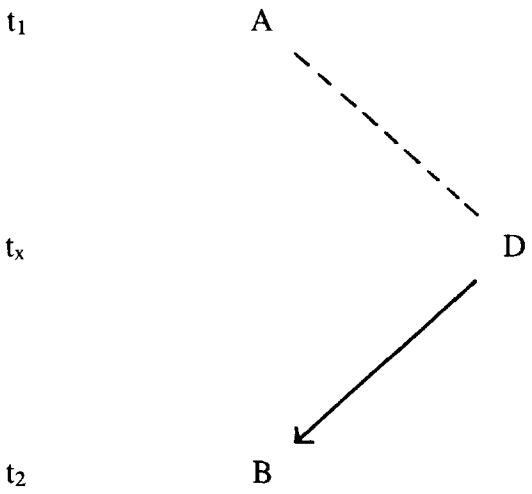

Figure 4.

such that each event is the direct cause of the next. Given any two events on a causal chain we can say that the earlier causes the later *indirectly*, via the intervening stages. In effect this defines causation (direct or indirect) ancestrally, in terms of direct causation: a cause is a direct cause or a cause of a cause.

The obvious objection to all this is that time isn't discrete, but dense. Between any two times there is always another. And this clearly invalidates the proposed approach. For, if we consider only the original discrete sequence of times, a factor A, at t_1, say, might appear to be a direct cause of B at t_2, even though it wasn't really a cause at all. Because even if it's not screened off from B by anything at t_1 or earlier, it might still be screened off from it by some D at t_x, halfway between t_1 and t_2, where D *isn't* in fact causally intermediate between A and B, but merely a confounding factor associated with A (because of some common causal ancestor, say). (See Figure 4: the solid arrow indicates causation, the dotted line probabilistic association.)

Well, we could deal with this case by considering a finer sequence of instants, which included all the times midway between the original times, and so included t_x. Then A would be exposed as not a genuine cause of B, for although D would count as a direct cause of B, A wouldn't be a direct cause of D.

But of course the difficulty would still lurk in the interstices between the half instants. But now the solution should be clear. What we need to consider is the infinite series, s_1, s_2, . . . , of finer and finer sequences of instants between t_0 and t_k. If A really isn't a genuine causal ancestor of B, then at some point in the series we will have a fine enough discrimination of instants for it to be exposed as such, in the way that A was exposed as an imposter by D above. Conversely,

if A *is* a genuine causal ancestor of B, then, however far we go down the series, the finer and finer divisions will all present it as a direct cause of a direct cause . . . of a direct cause of B.

Since time is dense there aren't, strictly speaking, any direct causes, and so, given the earlier definition, no indirect causes either. But that doesn't matter. We can regard the idea of direct and indirect causation as defined relative to a given fictional division of time into a discrete sequence of instants. And then we can define genuine causal ancestry as the limit of indirect causation in the infinite series of such fictional divisions, in the way indicated in the last paragraph.

15. Causal Asymmetry

I'm still cheating. I have now given up the earlier simplification that the only relevant times are "earlier" and "later," and explained how to deal with the fact that in between any two different times there are always infinitely many more. But the analysis still depended on a crucial implicit assumption about the relation between causation and time, namely, that the causal direction always lines up with the earlier-later direction.

In effect what I showed in the last section was that genuine causal connections between finitely separated events can be explained in terms of causal connections between, so to speak, infinitesimally separated events. But I simply took it for granted that when we had such an infinitesimal causal connection between some A and B, then it was the earlier A that was the cause of the later B, not vice versa.

I would rather not take this assumption of the temporal directionality of causation for granted. For one thing, there is nothing in the probabilities as such to justify the asymmetry: the relation of having a unscreenable-off probabilistic association is an entirely symmetric one. But that's not the crucial point. If nothing else were at issue, there wouldn't be anything specially wrong with reducing causation to probability and temporal direction, rather than to probability alone. But something else is at issue. There is a good independent reason for being dissatisfied with building temporal direction into the analysis of causation. Namely, that there are obvious attractions to the converse reduction, of temporal direction to causation. After all, what is the past, except those events that can affect the present, including our memories? And what is the future, except those events that the present, including our present actions, can affect? But if we want to expand these thoughts into an analysis, we'd better not build temporal direction into causal direction.

So the problem is to explain causal asymmetry without assuming that causes always precede their effects in time. There isn't any question of treating this problem fully here. But let me try to give some idea of an approach which makes use of some of the notions I have developed in the present paper.

This approach is defended in greater detail in my (1985b). In that paper I begin

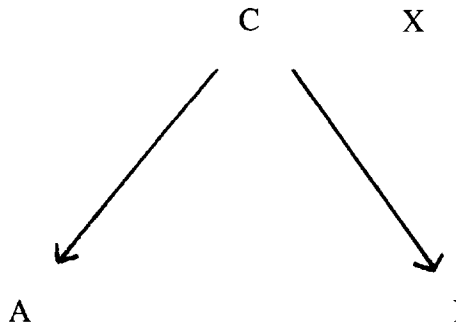

Figure 5.

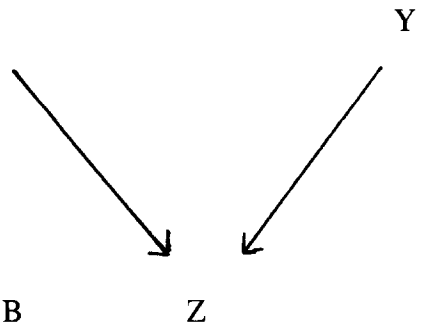

Figure 6.

with the fact that we often find that probabilistic associations between some A and B are screened off by a third factor C. I then observe that such cases are characteristically those where C is a common cause of A and B (or where C is causally intermediate between A and B). But I point out that we don't find this pattern when some Z, say, is a common *effect* of some X and Y.

The thought I pursue is that the probabilistic difference between Figures 5 and 6—screened-off associations in Figure 5, but none in Figure 6—is symptomatic of the differences in causal direction involved. This leads me to look for some independent explanation of the probabilistic differences, which might then serve as an analysis of causal direction.

The explanation I offer is that the screenable-off associations arise because (a) the probabilities involved are mixed, and (b) the mixing factors satisfy certain independence assumptions. Suppose the probability of A given C is a mixture: together with some background conditions C fixes a certain chance for A, together with others it fixes different chances. And suppose that the same is true of the probability of B given C. Then one can argue that, if the background conditions which together with C are relevant to the chance of A are *probabilistically independent* of those which together with C are relevant to the chance of B, then there will be a probabilistic association between A and B, and that association will be screened off by C.[5]

It follows that if, in Figure 6, there are sets of background conditions, together with which Z fixes chances respectively for X and Y, these background conditions can't be probabilistically independent, for if they were then there would be an X-Y association which was screened off by Z. And I confirm the analysis by showing that, in actual cases of joint causes X and Y of a common effect Z, the background factors required to specify laws which run, so to speak, from Z to the respective chances of X and Y, will manifestly not be probabilistically independent.

So I suggest the following account of causal direction. The properties whose

causal relationships we are interested in are generally related in "mixed" ways: the chances of one property given another will vary, depending on the presence or absence of various sets of background conditions. The causes can then be differentiated from the effects by the principle that the various sets of background conditions, together with which a cause is relevant to the chances of its various effects, are mutually probabilistically independent, whereas the converse principle got by interchanging "cause" and "effect" is not true.

On this suggestion, the directionality of causation doesn't lie in the structure of the lawlike connections between events themselves, so much as in the further probabilistic relationship between various sets of background conditions involved in such lawlike connections. It may seem odd to attribute causal direction not to the causal links themselves but to the satisfaction of probabilistic independence conditions by (often unknown) background conditions. It is worth noting, however, that quite analogous explanations can be given for two other puzzling physical asymmetries, namely the fact that entropy always increases, and the fact that radiation always expands outward. Although the underlying laws of physics permit the reverse processes to happen, when the laws of physics are combined with certain assumptions about the probabilistic independence of initial, as opposed to "final," micro-conditions, then the asymmetrical behavior can be derived.

It is also worth noting that the analysis of causal direction that I have outlined in this section is *not* committed to the "principle of the common cause": I am not assuming that for every correlation between spatio-temporally separated events there is some common cause that screens off their association. My claim is only that *if* there is such a screener-off, *then* it will be a common cause of its two joint effects, rather than a common effect of joint causes. Note in particular that there is nothing in this to conflict with the existence of unscreenable-off correlations, as in the EPR experiments. (After all, everybody agrees that the intuitive significance of such unscreen-offability is precisely that there couldn't be a common cause of the effects on the two wings.)

Now that the EPR phenomena have been mentioned, it will be worth digressing briefly and saying something more about them. Maybe the EPR phenomena don't causes any difficulties for my analysis of causal direction in terms of screening off. But, even so, they do raise a substantial problem for my overall argument. For they seem to provide direct counterexamples to the reductionist thesis itself.

In the EPR experiments the chance of a given result on one wing is increased by the chance of the corresponding result on the other wing, and this correlation isn't screened off by anything else. Given my overall reductionism, this ought to imply that there is a direct causal connection between the results on the two wings. But we don't want this—apart from anything else, such instantaneous action at a distance would seem to contradict special relativity.

However, I think that the analysis developed so far yields a natural way of ruling out the EPR correlations as causal connections. As a first step, let me add the

following requirement to the analysis so far: direct causal connections should be concatenable into causal chains: correlations not so concatenable should be disqualified as causal connections on that account. This might seem trivial: once we accept that A is a cause of B, then won't we automatically conclude, given any D that causes A, that we have a causal chain from D to B through A? But, trivial as it is, this requirement suffices to rule out the EPR correlations as causal connections. For Michael Redhead has shown that part of the weirdness of the EPR correlations is that they are *not* concatenable into causal chains (Redhead, forthcoming).

More precisely, Redhead shows that if A and B are correlated results on the two wings of an EPR experiment, and D is a cause of A, then A doesn't behave probabilistically like a link in a causal chain from D to B: A doesn't screen off the correlation between D and B. This is because, when A comes from D, the A-B correlation is itself altered, in such a way as to undermine the screening-off feature. As Redhead puts it, the A-B correlation is not *robust*, in that it is sensitive to factors which affect the presence or absence of A.

I am assuming here that it is essential to the existence of a causal chain that intermediate links should screen off earlier stages from later stages. I admit that nothing in the earlier discussion of causal chains guarantees this assumption. But it seems to me that it flows naturally from the arguments of the last two sections. Earlier in this section I suggested that it is constitutive of the common-cause-joint-effect relationship that common causes should screen off the correlations among their joint effects. So let me make an analogous suggestion about the causal chains introduced in the last section: namely, that it is constitutive of the idea of one factor being causally intermediate between two others that it should screen off the correlations between them.

The interesting question which remains is whether this last screening-off pattern can be reduced to independence requirements on background conditions, analogous to the suggested reduction of the common-cause-joint-effect pattern. I make some brief comments on this issue in my (1985b). More detailed investigation will have to wait for another occasion.

16. Digression on Independence Requirements

The last section involved certain independence assumptions about background conditions. In this section I would like to make some further points related to such independence assumptions. Most of this is about technical difficulties in my overall argument. Some readers might prefer to skip ahead to the next section.

Let us go back to the idea that mixed probabilities can be reliable guides to population causation. The danger with such probabilities was that they might be spurious, as well as mixed, in which case they would be misleading about population causation. My response was to point out that this threat could be blocked by

dividing the overall reference class into cells within which the putative cause C isn't probabilistically associated with any other relevant background factors, and seeing whether C still makes a difference to the probability of E within such cells.

It has been important to a number of my arguments that this doesn't necessarily require dividing the reference class into *homogeneous* cells. It is precisely because not *all* the other conditions relevant to E will in general be associated with C that we are ever able to reach conclusions about population causes from mixed probabilities. Moreover, this fact (or, more accurately, our believing this fact) is also a precondition of our acting rationally on probabilities that we believe to be mixed.

Nancy Cartwright has asked (in conversation) why we should suppose that, once a few confounding factors have been taken into explicit account, the remaining relevant conditions will generally be probabilistically independent of C. That is, why think that once we have made some fairly gross and inhomogeneous division of the reference class, all remaining factors will be independent of C within the resulting cells?

I don't have any basis for this presupposition, beyond a metaphysical conviction, which I may as well now make explicit. This is simply the conviction that in general different properties are probabilistically independent, except in those special cases when they are (as we say) causally connected, either by one being a causal ancestor of the other, or by their having a common causal ancestor.

I can't explain why the world should be like this. But I believe that it is, and, moreover, I believe that if it weren't it would be a very different place. If it didn't satisfy this general principle of probabilistic independence, we wouldn't be able to infer population causes from mixed probabilities, nor therefore would we be able to act on such probabilities. And indeed, if there is anything to the arguments of the last section, there wouldn't be any causal direction in such a world either.

There is, however, a difficulty which arises in connection with this independence principle, and which I rather slurred over in my (1985a). Consider the old chestnut of the falling barometer (B) and the rain (R). Suppose, for the sake of the argument, that rain is always determined when it occurs, either by a fall in atmospheric pressure (A) and high humidity (H), or by one of a disjunction of other factors, which I'll write as Y:

(1) A.H or Y $\leftrightarrow$ R.

Suppose also that A and X (the barometer is working) determine B, and so does Z (the kind of barometer malfunction which makes the barometer fall even though the atmospheric pressure hasn't):

(2) A.X or Z $\leftrightarrow$ B.

(Throughout this section I shall assume that all events have determining causes. Most of the arguments I give will be generalizable to indeterministic causes, or, equivalently, to the deterministic causation of chances.)

Now, if (1) and (2) are true, there is a sense in which falling barometers are, in Mackie's terminology, "inus conditions" of rain.

For it immediately follows from (1) and (2) that $B.-Z.H \rightarrow R$. This is because $B.-Z$ ensures A, by (2): if the barometer is not malfunctioning ($-Z$) and it falls (B), then the atmospheric pressure must have fallen. And so if we have high humidity (H) as well, it'll rain, by (1).

Moreover, we can no doubt cook up a Q which covers all the other causes of rain apart from drops in atmospheric pressure (Y), and also covers those cases where the barometer doesn't fall when the pressure falls in high humidity. Which will give us:

(3) $B.-Z.H \text{ or } Q \leftrightarrow R$.

Which is what I meant by the barometer being an "inus condition" of rain.

Equivalence (3) means that there is a sense in which $-Z$, H and Q are background conditions which, along with B, fix the chance of rain. Which means, given everything I've said so far, that once we've divided our reference class up enough to ensure that B is no longer associated with $-Z$, H and Q, we can draw conclusions about whether or not B is a population cause of R by seeing whether it is still probabilistically relevant to it.

The trouble, in this particular case, is that we won't ever be able to divide up our reference class in such a way as to get rid of confounding associations. For $-Z$ was specified as the absence of the kind of malfunction which makes the barometer fall, and clearly that's going to remain (negatively) associated with B, the barometer's falling, however much dividing up of the reference class we do.

But this is now somewhat paradoxical. For surely it is intuitively clear, quite apart from all these messy equivalences, that we *can* find out, from appropriate mixed probabilities, whether or not barometers cause rain. There is an initial probabilistic association between falling barometers and rain: falling barometers mean that rain is likely. But, by looking more closely at different kinds of cases, though without necessarily identifying all factors relevant to rain, it is in practice perfectly possible to show that this initial association is spurious.

Let me spell out the paradox. We know that we can expose the barometer-rain correlation as spurious without getting down to pure probabilities. But (3) and my general argument seem to imply that we oughtn't be able to do this, since we can't get rid of the confounding association with -Z without dividing the reference class *ad infinitum*. True, -Z is rather different from most of the confounding factors we have met so far, in that -Z is *negatively* associated with B, and so threatens to produce a spurious null or negative correlation between B and R, rather than the spurious positive correlation threatened by the usual kind of confounding factor. But the point remains. How can we be confident that the statistics show that B and R are *genuinely* null correlated, rather than only spuriously so, even though the negative confounding factor -Z hasn't been controlled for?

We need to take a few steps back to disentangle all this. The first thing to note is that, even if there's a sense in which barometers are "inus conditions" of rain, we certainly don't want to count them as causes of rain on that account. Barometers don't cause rain. The moral of the equivalences (1)-(4) isn't that barometers cause rain, but simply that inus conditionship isn't enough for causation.

It is obvious enough how we might strengthen the idea of inus conditionship to rule out cases like the barometer. The barometer gets to be an inus condition only by proxy, so to speak: as the derivation of (3) makes clear, it only suffices for rain by virtue of the fact that the atmospheric pressure has already always fallen on the relevant occasions. What is more, the atmospheric pressure sometimes suffices for rain on occasions when the barometer doesn't fall. In my (1978a) and (1978b) I said that in such situations the atmospheric pressure *eclipses* the barometer as an inus condition of rain. And I hypothesized that in general causation required *uneclipsed* inus conditionship, rather than inus conditionship alone.[6]

The importance of the suggestion that causation is equivalent specifically to *uneclipsed* inus conditionship is that it enables us to make a rather stronger claim about the connection between probabilities and population causation than anything we've had so far. In a number of places I've assumed that for any putative cause C and effect E there will be background conditions X which together with C ensure E, and also that there will be other sets of factors, Y, which don't include C, which also ensure E: C.X or Y $\leftrightarrow$ E. And then I've claimed that, once we've divided up our reference class into cells within which C is no longer associated with X or Y, then we'll get remaining correlations between C and E if and only if C is a population cause of Y.

But if causes are uneclipsed inus conditions we can say something stronger. Namely, that as soon as we've divided our reference class into cells in which C is no longer *associated with Y*, we'll get remaining correlations if and only if C is a population cause of E. The point here is that with uneclipsed inus conditionship we don't any longer have to control for X. As I show in my (1978a) and (1978b), the condition required for remaining correlations to be necessary and sufficient for uneclipsed inus conditionship is only that C not be associated with the other, independent factors ensuring E – it doesn't matter if C is associated with the factors that it acts in concert with.

Let us return to the barometer. The puzzle I raised was that, while in practice we can clearly use mixed probabilities to find out whether or not barometers cause rain, the condition required for the drawing of such inferences seems not to be satisfiable. For barometer falls will continue to be (negatively) associated with -Z, however much dividing up of the reference class we do.

But we now see that the requirement that any association with $-Z$ be eliminated is too strong. Given that not all inus conditions are causes, but only uneclipsed ones, we can relax the requirement on eliminating confounding factors.

We are in a position to draw conclusions about the barometer as long as we have assured ourselves that it is not associated with any of the other, independent sources of rain. And that we can easily do, by looking separately at classes of cases where the barometer falls, but the atmospheric pressure doesn't, and vice versa. Within these classes we will find that there are no remaining correlations between barometers and rain. And this then entitles us to conclude that barometers don't cause rain.

If what we wanted were conclusions about inus conditionship simpliciter, then we *would* need to control for $-Z$. And so we wouldn't be able to decide that the barometer wasn't an inus condition from the null partial correlations appealed to in the last paragraph, because those correlations are in classes where F is still negatively associated with $-Z$. But that's just as well, since the barometer *is* an inus condition of rain, albeit an eclipsed one. Indeed this shows, so to speak, why the association with $-Z$ has to be ineliminable, given that F has null partial correlations with R: for if the association with $-Z$ were ever eliminated, then we would be entitled to infer, from the null correlations, the false conclusion that F isn't an inus condition of R.

17. Causal Processes and Pseudo-Processes

Some recent writers working in the probabilistic tradition have turned against the reductionist idea that causal relationships can be defined in terms of probabilistic relationships between properties. Instead they suggest that causation is primarily a matter of *causal processes*, and that generic causal relationships between properties need to be defined in terms of causal processes, rather than conversely. Thus Wesley Salmon says, "The basic causal mechanism, in my opinion, is a causal process that carries with it probability distributions for various kinds of interactions" (1984, 203). And Elliott Sober says, "Connecting processes exist independently of the events they connect. Such processes are like channels in which information flows: the existence of the channel does not imply what information (if any) actually flows over it" (1986, 111).

In this section I want to show that, while the notion of a causal process is indeed an important notion, it can be defined straightforwardly enough in terms of the concepts already developed in this paper, and does nothing to suggest those concepts are inadequate for a reductionist theory of causation.

Let me begin with Salmon's distinction between causal processes and pseudo-processes (1984, 141–42). A causal process can transmit marks. Pseudo-processes cannot. So a moving shadow on a wall is a pseudo-process, for you can't alter the later characteristics of a shadow by operating on its earlier stages. The expansion of radiation, or the persistence of a normal physical object, on the other hand, are causal processes, since the later stages of such processes can generally be altered by acting on their earlier stages.

This is indeed a significant distinction. But it would be a mistake to think that it is somehow primitive to the theory of causation. Suppose we call any sequence of space-time points a space-time *worm*. Some space-time worms will be distinguished by the fact that some of the properties possessed by earlier stages will be correlated in various ways with some of the properties possessed by later stages. Let us call these worms *processes*. Processes are those worms which carry sequences of correlated properties.

So far this goes for causal and pseudo-processes alike. The difference is that the sequences of properties carried by *causal* processes comprise causal chains, in the sense outlined in Section 14 above: insofar as the probabilistic relevance of earlier properties in a given sequence to later ones is screened off, it will only be by properties intermediate in that sequence. The sequences of properties carried by pseudo-processes, on the other hand, are not causal chains: for the probabilistic relevance in such sequences will often be screened off by facts which are not part of the pseudo-process at all.

To illustrate, think of people as causal processes. Many of the earlier properties of people are relevant to their later properties. Childhood weight is relevant to adult weight, childhood hair color to adult hair color, etc. And in these cases nothing fully screens off adult weight (hair color, etc.) from childhood weight (hair color, etc.) except the intermediate weights (etc.). But in the case of a moving shadow, say, the intermediate shape of the object casting the shadow will screen off the shadow's later shape from its earlier shape: which stops the sequence of shapes of the shadow being a causal chain, since the intermediate shape of the object casting the shadow isn't itself part of that sequence.

So we can define causal *processes* in terms of causal *chains*. Causal processes are space-time worms that carry ensembles of causal chains. If we like, we can think of causal processes as bundles of causal chains.

There is nothing in anything said so far to motivate any revision of my overall reductionist thesis. However, both Salmon and Sober seem to think that once we have a notion of causal process, then this allows for a new kind of causation. They hold the later features of a causal process should be deemed to be caused by the relevant earlier features, even in cases where our reductionist notions would give the opposite answer. Thus, to pursue the above example, they would say that if a fat child grows into a thin adult, the adult thinness is caused by the childhood fatness.

The idea here seems to be that, since weight in humans is a property carried by a causal process, the later weights of humans are always caused by their earlier weights. But there is nothing in the notion of causal process, as I have explicated it, to warrant this. Causal processes are picked out as such because they carry causal chains. Causal chains are a matter of certain earlier properties making certain later ones more likely than they would otherwise be. Nothing in this requires us to say that when, in a particular case, the relevant later property doesn't occur,

that its noninstantiation was caused by the earlier property. We can still insist that an earlier property only affects another if it makes its chance higher than it would otherwise have been.

18. Negative Causes

In this final section I want to look more closely at the possibility of "negative causes," causes that make their effects less likely. In the last section I argued that the notion of a causal process as such yields no argument for such negative causes. But in a sense this puts the cart before the horse. For perhaps there are independent arguments for admitting negative causes. And if there are, then this will itself provide a reason for wanting causal processes as a primitive component in the theory of causation, over and above any arguments considered in the last section.

The reason would be this. As long as we assume that all causes are positive, then we can plausibly define causation in terms of one event fixing an increased chance for another, and we can define causal processes in terms of causal chains as in the last section. But if causes can either increase or decrease the chances of their effects, then it seems highly unlikely that probabilistic relationships alone will distinguish cause-effect pairs from others – for all kinds of events either increase *or* decrease the chances of other events. (See Sober 1986, 99.) So if we allow negative causes, it seems that we will need some independent notion, such as that of a causal *process*, to provide the cement in the causal relationship, that is, to tell us which pairs of (positively or negatively) probabilistically related events are in fact related as cause and effect.

This then is an argument *from* negative causes *to* causal processes. But, still, why should anybody accept negative causes in the first place? If somebody grows up thin after being a fat child, it seems quite counterintuitive to say that they are now thin *because* they were once fat. The natural thing to say, surely, is that they are now thin *despite* once being fat.

Even so, a surprising number of philosophers working on probabilistic causation have been prepared to override this intuition and admit negative causes. To understand this tendency we need to recognize an implicit theoretical rationale, which goes back to the original reasons for switching from the traditional deterministic theory of causation to the S-R view. One initial attraction of the S-R view of causation was that it allowed us to go on having causation in the face of indeterminism. However, this initial attraction loses some of its force once we attend to the fact that in general causes produce their effects only via chains of intermediate events. For it always seems possible that one or more such intermediate events might reduce the chance of the final result, not increase it. If we allow such intermediate negative causes to break the overall chain, then it seems that we won't have much causation in an indeterministic world after all. So to preserve the at-

tractions of the S-R over the traditional deterministic view, we had better allow negative causes in a causal chain to count as intermediate causes. (See Salmon 1984, ch. 7, where this line of thought is clearly implicit.)

The kind of example at issue was first discussed in Suppes's original (1970) formulation of the S-R view. A man hits a golf ball. In its flight it strikes a tree. It lands in the hole. Golf balls that hit trees are less likely to go in holes than those that don't. But, still, wasn't the ball caused to go in the hole by hitting the tree?

A natural initial reaction here is to say that given a full enough description of the golf ball's initial path and the precise way it hit the tree, the ball's hitting the tree would no doubt have made it more likely to go in the hole, not less. But the example can be tightened up. The initial shot is a good one. But a squirrel runs out onto the fairway and kicks the ball as it goes past. The kick, when described in full detail, yields an objective probability distribution over subsequent trajectories for the ball which make it far less likely than before that the ball will go in the hole. But the ball does go in the hole. Didn't the squirrel's kick cause it to go in the hole? (See Sober 1986, 99.)

But now that the case has been tightened up to make it clear that the kick objectively *reduced* the chance of the result, then it seems to me as before that it is highly counterintuitive to say that the kick caused the ball to go in the hole. True, the ball ended up in the hole. But that was just a matter of luck, given the kick. It wasn't because of it.

If the price of defending the S-R view over the traditional view is admitting negative causes, then so much the worse for the S-R view. Surely we would be better off simply resting with the traditional view, which didn't offer to transmit causation across indeterministic gaps in the first place. At least that would remove the temptation to adopt the curious view that results can be caused by events that lower their chances.

What is more, defending the S-R view by allowing in negative causes reduces the power of the S-R view to explicate causation in the first place. For, as explained at the beginning of this section, once we admit negative causes, then we will also need to introduce some independent notion of a causal process to distinguish causal connections from others. So in order to defend the S-R view via negative causes it will be necessary to relegate it to a relatively minor role in the theory of causation.

I don't want to suggest that all this yields a conclusive argument against the S-R view. For it seems to me that the original rationale for admitting negative causes is weak, even from the perspective of the S-R view. That is, I think that S-R theorists can perfectly well accommodate the squirrel-type cases without resorting to negative causes.

There are a number of alternative moves open to S-R theorists here. For instance, they could hold that, whenever an earlier event makes a later one more likely, and the later event happens, then the earlier event causes the later event,

even if an intermediate event along the way later shifted the probability of the eventual result downward to some degree (and so isn't itself to be deemed a cause). But perhaps this raises difficulties of its own. If the probability-lowering intermediate event isn't a cause, then we will be committed to a kind of causal action at a temporal distance, for the initial event will cause the final event without the causation being continuously transmitted in between. Such causal action at a distance seems unattractive.

A better move for the S-R theorist seems simply to give up the squirrel-type cases as causal chains, and allow that, if there are probability-lowering intermediaries, then the causal chain is broken. This will mean that S-R theorists will have to accept that there is rather less causation in the world than we used to suppose before quantum mechanics. But there is no reason why the S-R view should be committed to preserving universal causation.

Exactly how many *prima facie* causal chains would in fact be broken by probability-lowering stages is an interesting question, which seems to me to deserve rather more consideration than it has received in the literature. For it by no means immediately follows from quantum mechanics that every sequence of events contains some probability-lowering stage.

Even if it were concluded that *most* apparent causal chains get broken by probability-lowering stages, the S-R theory would scarcely be worse off than the traditional deterministic view on this score. For the traditional view will regard causal chains as broken by *any* indeterministic stages, whether or not they lower the probability of the eventual result. The traditional view may well still have *some* causal chains. (Quantum mechanics doesn't imply that every sequence of events contains indeterministic stages, any more than it implies that every sequence contains probability-lowering stages. It yields indeterminism only where wave functions "collapse" in "measurements," and it's an open question exactly which events should be so conceived.) But the traditional view certainly won't have *more* causal chains than the S-R view.

Doesn't this now yield a kind of quantitative argument for the S-R view over the deterministic view? The S-R view preserves more *prima facie* causal chains than the deterministic view. But why suppose that the maximal preservation of *prima facie* causal chains as real causal chains is a desideratum in the theory of causation? Some S-R theorists seem to think that it is, and indeed it is this preservationist concern that moves them to countenance negative causes. But I have argued that negative causes are a bad idea, and that the best way to do without them is to recognize that certain sequences of events that we took to be causal before quantum mechanics might well not be causal. Given this, it is scarcely decisive against the traditional view of causation to say that it denies continuous causation to all sequences involving indeterministic stages. For the traditional theorist can counter by pointing out that the S-R view also denies continuous causation to certainly apparently causal sequences—namely, sequences like the squirrel case

where an intermediate event lowers the chance of the eventual result. At which point we can discuss exactly which class of *prima facie* cases really ought to be deemed to lack continuous causation. All the old arguments will come back into play, and once more we will be led to a stand-off.

In developing the ideas in this paper I have benefited from a long history of discussions with Nancy Cartwright and Hugh Mellor. An early version of the paper was delivered in 1985 at a seminar at the Minnesota Center for the Philosophy of Science in the program run by Wesley Salmon and Philip Kitcher; I would like to thank both of them for inviting me, and for much helpful discussion. Some of the later parts of the paper emerged from a symposium between Elliott Sober and myself, chaired by Hugh Mellor, at the joint session of the Aristotelian Society and Mind Association in 1986, and were improved by what they said there. I would also like to thank Paul Humphreys for discussing a draft of the paper with me.

Notes

1. To avoid confusion I should say that I regard *chance* as the fundamental notion, rather than long-run relative frequency in a reference class, or anything along such lines. Probabilities are always chances, or chance combinations of chances. A reference class should be thought of as a construction out of its homogeneous subdivisions: the probability in an inhomogeneous cell C is then the weighted average of the chances in the homogeneous subdivisions of that cell, with the weighting factors as the probabilities, given C, of being in each such homogeneous subdivision.

2. Perhaps I should make it clear that in talking about "single cases" and "single case causation" I am not endorsing the view that causation needs to be analyzed in terms of some irreducible relation between particular events. I still intend to show that causation reduces to generic (and not explicitly causal) relations of association between properties (together, obviously, with particular, but nonrelational, facts about which properties are instantiated on which occasions). The point of talking about single cases is simply to make it clear that the generic associations relevant in any particular case are those involving *all* probabilistically relevant properties present on that occasion, and not any mixed-probability-yielding subset thereof.

3. If determinism is true, how can there be any probabilities other than nought and one? But note that the probabilities in question need only the *initial conditions*, X or Y, to have probabilities different from nought and one, given S. And that's certainly consistent with C always being determined, in the sense of there always being complete determinism, so to speak, *between* S, X and Y, on the one hand, and C, on the other. But still, the objection might be pressed, how can X and Y have real probabilities different from nought and one, if everything's always determined? I used to think that this problem could be pushed back to infinity, with probabilities always deriving from probability distributions over prior initial conditions, which themselves were determined by prior initial conditions, etc. But now I'm not so sure. Unless there were some real chance events in the background somewhere, I don't think the probabilities over the initial conditions would display the kind of randomness and stable long-run frequencies characteristic of real probabilities. But we can by-pass this worry: for the purposes of this paper let us understand determinism not as the cosmic claim that there are no chances anywhere, but as the context-relative thesis that the events under consideration are always determined by their recent local histories.

4. In fact the two definitions aren't quite equivalent. Imagine a case where the chance of a given result is overdetermined, in the sense that an alternative chain of events, not present in the actual circumstances, would have come into play and ensured the same chance for E if C had been absent. In this case it is true that C and the other factors actually present fix a probability-increasing reference

class for E (and correspondingly E does intuitively depend causally on C). Yet it's not true that E would have had a lower chance if C had been absent. This case shows that to preserve the counterfactual formulation of causal dependence, we need to qualify it so as to make it sensitive to the result's actual causal ancestry in the circumstances. Though I shall not explicitly discuss this issue any further, the discussion of "causal chains" in section 14 provides the materials for the necessary qualification. See David Lewis's treatment of this issue in his (1973).

5. This conclusion now seems to me somewhat less compelling than I supposed in my (1985b). In that article I concentrated on the deterministic case where C plus background factors (X, say) determines A, and C plus Y similarly determines B. And in that case it is certainly true that an appropriate probabilistic independence condition on X and Y will force A and B to have a probabilistic association that is screened off by C. But if C.X only determines that A should have some chance p, and similarly C.Y only determines that B should have chance q, where p and q are less than one, then there is still room, so to speak, for A and B to coordinate themselves in such a way as to prevent screening off by A. I am thinking here of the kind of phenomenon observed in EPR situations. But, even so, the basic idea that screening off comes from mixtures still seems to me sound. For even in the indeterministic case the independence of X and Y will guarantee that A's *having chance* p will be probabilistically independent within C of B's *having chance* q, and to that extent at least we will tend to get screening off of A-B correlations by C.

6. The claim that causation is specifically uneclipsed inus conditionship faces its own difficulties. Suppose an intermediate cause, D, of some result, E, always comes from a given prior cause, C. And suppose also that there's an alternative route from C to E, that is, that C sometimes causes E via a chain that doesn't include D. Then C eclipses D as a cause of E. But, by hypothesis, D does sometimes cause E. Here I am inclined to appeal to our intuition in such cases that if D *were* ever to be present without C, then E would still occur. (Compare the contrasting intuition, that even if a properly working barometer were ever to fall without a drop in atmospheric pressure, that wouldn't produce rain.) In terms of chances rather than counterfactuals, I am suggesting that we should look to the cell of the reference class defined by D & not-C to see that D rather than C is on occasion the direct cause of E (and thus to see that D isn't "really" eclipsed by C). The remaining problem, of course, is that there aren't any actual cases in the relevant "D & not-C" cell: I am invoking a probabilistic law without any this-worldly instances. But I'm not too worried. Quite apart from the present problem, there is clearly a general need for some philosophical account of uninstantiated laws. I don't want to belittle the philosophical difficulties involved in giving such an account. But since they will have to be dealt with anyway, there's no reason not to appeal to uninstantiated laws to get out of the present problem.

References

Cartwright, N. 1979. Causal Laws and Effective Strategies, *Noûs*, 13:419–37.

Eells, E. 1982. *Rational Decision and Causality*. Cambridge: the University Press.

——, and Sober, E. 1983. Probabilistic Causality and the Question of Transitivity, *Philosophy of Science*, 50:35–57.

——. 1986. Common Causes and Decision Theory, *Philosophy of Science*, 53:223–45.

Jeffrey, R. 1983. *The Logic of Decision*, 2nd ed. Chicago: University of Chicago Press.

Kitcher, P. 1985. Two Approaches to Explanation, *Journal of Philosophy*, 82:632–39.

Lewis, D. 1973. Causation, *Journal of Philosophy*, 70:556–67.

——. 1981. Causal Decision Theory, *Australasian Journal of Philosophy*, 59:5–30.

Mellor, D. forthcoming. Indeterministic Causation.

Nozick, R. 1969. Newcomb's Problem and Two Principles of Choice. In *Essays in Honour of Carl G. Hempel*, ed. N. Rescher. Dordrecht: D. Reidel, 114–46.

Papineau, D. 1978a. *For Science in the Social Sciences*. New York: Macmillan.

——. 1978b. Salmon, Statistics and Backwards Causation. In *PSA 1978*, eds. P. Asquith and I. Hacking vol. 1, 302–13.

——. 1985a. Probabilities and Causes, *Journal of Philosophy*, 82:57–74.

——. 1985b. Causal Asymmetry, *British Journal for the Philosophy of Science*, 36:273–89.

Redhead, M. Forthcoming. Relativity and Quantum Mechanics—Conflict or Peaceful Coexistence, *Annals of the New York Academy of Sciences*.

Salmon, W. 1970. Statistical Explanation. In Salmon, Jeffrey, and Greeno.
——. 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
——. Jeffrey, R., and Greeno, J. 1970. *Statistical Explanation and Statistical Relevance*. Pittsburgh: University of Pittsburgh Press.
Sober, E. 1986. Causal Factors, Causal Inference, Causal Explanation, *Aristotelian Society Supplementary LX*:97–113.
Suppes, P. 1970. *A Probabilistic Theory of Causality*. Amsterdam: North Holland.