ABSTRACT. Practical causal generalizations are of the form \( A \) causes \( B \), where \( A \)'s are actions, as in 'Smoking causes cancer', and what makes these generalizations practical is that they guide persons in deciding whether to perform the actions they apply to. Though not all causal generalizations are equivalent to chance-inequalities of the form 'the chance of \( B \) given \( A \) is greater than the chance of \( B \) without \( A \)', it is hypothesized that this is always valid in the case of practical generalizations. This would imply that, without constraints such as temporal order or directness, Suppes's *prima facie condition of causality* captures the entire content of a causal generalization of the kind we are concerned with. However, this presupposes a concept of *chance* that is neither subjective nor statistical, and unsystematic remarks are made on how these chances should be estimated if decisions based on them are to have the results that agents desire. Other topics touched on include the concept of a *voluntary act* as distinct from an 'event' that may happen independently of the will.

PART I. INFORMAL CONSIDERATIONS

1. Introduction

'Smoking causes cancer' and 'Taking quinine wards off malaria attacks' are propositions of the sort that this paper will be about. These are general and not particular statements, they are causal, and right or wrong they guide our actions – which is explained by the fact that the 'causes' they speak of, like smoking or taking quinine, are actions. They are things that we can do, not just 'events' like eclipses of the sun that happen independently of our wills. Of course they are vague and subject to huge numbers of exceptions, but they are the common coin of practical wisdom, and much of our practical education consists in acquiring this wisdom. As such I think it is important to understand this guidance and to evaluate it.

The remaining sections of Part I will discuss this matter informally, and Part II will describe elements of a mathematical formulation. We will begin with a comment on connections between correlations and causality, which will be discussed in greater detail later.
2. Causes and Correlations

It used to seem to me that whether or not correlations establish a ‘real’ causal connection between smoking and cancer, they are the only things that should matter for practical purposes. That is because the practical use to which we put the information conveyed by ‘Smoking causes cancer’ is in making decisions, in particular as to whether to smoke, and to make these decisions all we need to know is what will happen or what is likely to happen if we smoke. From that point of view it seemed to me that it shouldn’t matter whether smoking causes these things to happen. Experiences – smoking, having cancer, etc. – are the practically important things, and what produces them – causal connections, predestination, God’s will, or whatever – can seem to be only matters of metaphysical interpretation without intrinsic utility. Therefore I thought that science was practically justified in concentrating on correlations per se, since it seemed to be immaterial whether they were real or spurious.

But the example of ‘Taking quinine wards off malaria attacks’ shows that that can’t be right.

3. The Practical Importance of Casuality

‘Taking quinine wards off malaria attacks’ affirms that quinine taking has a negative causal influence on attacks of malaria, and if for practical purposes, this were equivalent to a statistical claim then there should be a negative correlation between quinine taking and malaria attacks. But very probably the correlation between them is positive, because persons who take quinine tend to be malaria sufferers, and though taking quinine may reduce their chances of having malaria attacks, they are nevertheless more likely to have them than most persons who don’t take quinine, most of whom aren’t malaria sufferers. One way of putting this is to say that although taking quinine may ward off a malaria attack, it is a sign that one may be immanent. But, assuming that persons take quinine to ward off malaria attacks, they are guided by what taking quinine causes, or influences, and not by what it is a sign of. If that is so, then the causal content of a maxim like ‘taking quinine wards off malaria attacks’ is essential in guiding our actions.

But this creates a problem for empirical science, because observation is restricted to discovering empirical correlations, which show what things are signs of and not necessarily what they cause. Hence, it would seem that something essential to the practical ‘application’ of claims
like ‘Smoking causes cancer’ and ‘Taking quinine wards off malaria attacks’ transcends experience, and lies beyond the reach of scientific evaluation. In a sense this generalizes Fisher’s criticism of the use of statistics on smoking and cancer to establish a causal link between them, by implying that no empirical statistics can do this. Of course this only reiterates Humean skepticism about the experiential status of causality, with the further remark that it matters.

But we shouldn’t give in to transcendentalism too easily, and now I want to comment on ways in which probabilistic formulations like the ones proposed in Patrick Suppes’s very important and influential book *A Probabilistic Theory of Causality* (Suppes, 1970), may contribute, by reformulating causal claims in such a way as to bring them closer to empirical evaluation. I will concentrate on ‘Smoking causes cancer’, and approach a probabilistic formulation in two stages.

4. A Chance-Influence Formulation

The first step is to paraphrase ‘Smoking causes cancer’ as a *chance-influence claim*: ‘Smoking increases a person’s chances of getting cancer’. It seems to me that this is an improvement on ‘Smoking causes cancer’, because it allows for the fact that smoking doesn’t infallibly result in getting cancer – though it is still a good idea to avoid smoking because it increases the likelihood that a person will get cancer. No doubt something like this is the import of the Surgeon General’s warning ‘Smoking may be hazardous to your health’, which aims to discourage smoking without flatly asserting that it will result in harm.

But the chance-influence formulation is still causal, although it says that the immediate effect of smoking is not directly on cancer, but rather on the chances of getting it. So, even if this formulation is an improvement causality is still in the picture. The next move seems to get rid of it entirely, and to reduce everything to mathematical relations of chances.

5. A Chance-Inequality Formulation

This is to replace ‘Smoking increases a person’s chances of getting cancer’ by a *chance-inequality*, namely ‘A person’s chances of getting cancer are greater if she smokes than if she doesn’t’. This is something like the *prima facie* condition of causality that is focused on in *A Probabilistic Theory of Causality* (Suppes, 1970, p. 12), and one thing that
is important about it is that it doesn’t refer to causes or influences at all, even of chances. Instead, it states a mathematical relation between people’s chances of getting cancer if they smoke and their chances of getting it if they don’t. Of course going from causes to chances moves us from one transcendental realm to another, but this is still an important advance because the chances involved are generally assumed to be conditional probabilities, whose mathematical properties are well understood, in contrast to causal relations which are much more problematic from the mathematical standpoint.

Of course, how the chances are measured is a problem, but I think it is an advantage that they don’t necessarily have to be interpreted as relative frequencies. For instance, to go back to the malaria and quinine case, a person’s chances of suffering a malaria attack following taking quinine don’t have to be measured by the frequency of malaria attacks among persons who take quinine. That isn’t to say that the two are unrelated, but the looseness of the connection between them is important from the point of view of guidance for action, because that is more plausibly related to chances than to statistics. This will be returned to in Sections 8 and 9.

First, however, we must note a possible limitation of the inequality formulation that Brian Skyrms reminded me of, which suggests that no matter how chances are interpreted this formulation may only be valid for practical causal generalizations.

6. A Difficulty

We may take it for granted that there is a greater chance of rain when the barometer is low than when it is high, but that while a low barometer may be a sign of rain it doesn’t cause it. Given this we would make the negative causal claim, that a low barometer doesn’t cause rain, which could be restated as a negative chance-influence claim that a falling barometer doesn’t bring about an increase in the chance of rain. But that couldn’t be paraphrased as a chance-inequality, to the effect that the chance of rain is not higher if the barometer is low than if it is high. In short, while the crude causal generalization can be restated as a chance-influence, its replacement by an inequality doesn’t seem to be valid. But now I want to hazard a conjecture that could be quite important if it is true.
As far as I can see, the problem of the non-replaceability of chance-influences by chance-inequalities only arises in the case of non-practical causal generalizations, when the ‘conditions’ of the chances in the inequalities are things that happen or ‘are the case’, and not when they are things that are done. Thus, that a barometer is low is something that is the case, and the corresponding chance-inequality doesn’t correspond to a chance-influence. But smoking and taking quinine are things that are done, and in those cases chance-influences seem to be equivalent to inequalities. It is worth noting too that the rule that in the practical case chance-influences are equivalent to inequalities is borne out if ‘the barometer is low’ is replaced by ‘the barometer is lowered’ (i.e., persons lower the levels of mercury in them), which is something that it is possible to do. That lowering a barometer doesn’t increase the chance of rain is equally well expressed by saying that the chances of rain are not greater if a barometer is lowered than if it isn’t.\(^1\)

It is even possible to argue that chance influences ought to be replaceable by chance inequalities in the practical case. Take the malaria and quinine case, and suppose that a person wants to bring it about that her chances of having a malaria attack are as small as possible. If she can take the quinine, and she is aware that her chances of having an attack will be lower if she takes it than if she doesn’t, she can ‘apply’ this information by taking it. In other words, when the causes or ‘conditions’ involved in a chance-inequality are ones that persons can bring about, merely stating the inequality conveys the same practical information as the causal generalization because it tells people which condition to ‘realize’.

Of course we haven’t completely reduced practical causal generalizations to mathematical relations of chances, since we still have to explain the difference between things that are done and things that merely happen, and brief comments on that will be made in Section 10. But before that something should be said about estimation, and to lead into that we will return briefly to a theme that was introduced in Section 2. If ‘experiences’ like having cancer or suffering malaria attacks are the things that have intrinsic utility, why should their transcendental causes or chances matter?
7. The Practical Importance of Chances

Chances can’t harm a person, so why should he or she care what the chances of a malaria attack are, so long as the attack itself is avoided? In fact, it seems obvious that if a person could know whether or not taking quinine would be followed by a malaria attack, there would be no need to consider chances. Now, I suggest that considerations of chance enter the picture when this kind of knowledge isn’t available. When persons can’t know for sure at the time of acting what the consequences of their actions will be, it may be prudent to adopt policies that, while not always yielding desired results, do so ‘on balance’. If they can’t always avoid malaria attacks they at least have practical reasons for minimizing their frequency, and this seems to me to be what connects chances with practical values. This can be clarified a bit by bringing in Laws of Large Numbers.

Roughly, these imply that it is highly probable that there will be a very small difference between the frequency of malaria attacks that a person will suffer on a series of occasions and her mean chances of suffering them on those occasions, and therefore she will almost certainly minimize the frequency of her attacks if she minimizes her mean chances of suffering them. And, she can influence these chances by her actions, such as by taking quinine. This can be fine-tuned to take into account changing chances, like changes in the severity of a person’s case of malaria, and changing and competing values, such as her dislike of quinine and the cost and trouble of taking it. Then she can aim to maximize a time-weighted average of the values that interest her, calculated according to some expected utility formula, as will be discussed briefly in Section 13. In any case, however, she will want to compare the mean chances of the consequences of different possible courses of action, which leads to the problem of estimating these chances. This can be pictured in terms of an ‘urn analogy’.

8. Urn Analogy

As in the previous section, we will concentrate on the malaria and quinine case, though it is hoped that the transfer to other cases, such as that of smoking and cancer, will be obvious. In the malaria and quinine case we imagine persons choosing repeatedly between urns that are labelled ‘take quinine’ and ‘no quinine’, and after choosing an urn
they draw a ball at random from it, where all of the balls in the urns are marked ‘malaria attack’ or ‘no attack’.

Choosing the ‘take quinine’ urn corresponds to taking quinine, choosing the ‘no quinine’ urn corresponds to not taking it, drawing a ‘malaria attack’ ball corresponds to having a malaria attack, and drawing a ‘no attack’ ball corresponds to not having one. The proportion of ‘malaria attack’ balls in the ‘take quinine’ urn corresponds to the chance of having a malaria attack if quinine is taken and the proportion in the ‘no quinine’ urn corresponds to the chance of having one if quinine is not taken, and the latter is significantly greater than the former. But the analogy may be extended.

We will suppose that the urns from which persons choose and the proportions of the two kinds of balls in them can change not only from person to person but from occasion to occasion – so that the same person may draw from urns with different proportions of ‘malaria attack’ balls in them on different occasions. In particular, after a person contracts malaria the proportions of ‘malaria attack’ balls in the urns that he or she has to choose from will increase from 0 to some positive amount, though that is presumably less in the ‘quinine’ urn than in the ‘no quinine’ urn. Therefore what is pictured above represents the situation of the person who has malaria.

Assuming that the person wants to minimize the frequency of ‘malaria attack’ balls that she draws, laws of large numbers imply that she will be very likely to accomplish this by minimizing the mean proportions of these balls in the urns that she draws from. Therefore she should try to draw from urns with the smallest proportions of malaria attack balls in them, and the problem is to guess which those are. That is where estimation comes in, and now we will note a couple of things about how this appears from the point of view of the urn analogy.

9. Aspects of Estimation

One is that there are two ways of estimating proportions of balls with different markings inside urns, which I would call internal and external. The internal method is simply to look inside the urns. This would be something like directly studying the mechanisms by which quinine-ingestion affects the onset of malaria attacks, and these may be the best methods. But they may not always be feasible, and when they aren’t, external methods can be used in which, in the urn analogy, proportions
of different kinds of balls in urns may be estimated from the frequencies with which these kinds of balls are drawn from them. But this can be tricky.

What makes it tricky is the fact that what persons want to estimate are proportions of balls of different kinds in the urns from which they are going to choose, but what they have to base the estimates on are frequencies with which balls of these kinds have been drawn by other people from other urns, and that can be very unreliable. That’s what’s wrong in the quinine and malaria case. Judging one’s own chances of having a malaria attack by the frequency with which quinine-takers in general have had such attacks is to act as though one’s own chances are the same as those of typical quinine takers, while judging one’s chances of having an attack if quinine isn’t taken by the frequency with which persons who don’t take quinine have attacks is to act as though one’s chances were the same as those of typical non-takers of quinine.

Of course, the way to deal with this difficulty would be to compare the frequency of malaria attacks suffered by persons in ‘test groups’ to whom quinine is given with the frequency of attacks suffered by persons in ‘control groups’ from whom it is withheld, which would be analogous to forcing certain individuals to choose ‘take quinine’ urns and forcing certain others to choose ‘no quinine’ urns and then comparing the frequencies of ‘malaria attack’ balls drawn by the persons in the two groups. Given certain assumptions, this would resolve the problem, but if the persons in the two groups no longer acted voluntarily in choosing urns, what would be tested would no longer be a practical causal generalization, and we will end this part with a few final comments on this.

10. Actions and Events

If the conjecture formulated in Section 6 is right, that practical causal generalizations are equivalent to inequalities of conditional chances, the ‘causality’ that remains is confined to the conditions of these chances – that is, to what makes smoking and quinine-taking things that are done, as against things that are merely experienced. In effect, the inquiry shifts from the causal relation to the causal object, which, in the case of practical generalizations, is an action.
Now we must ask: beyond signalling that a chance inequality is practical and may be acted upon, is there a ‘real’, causal difference between, say, taking quinine intentionally and swallowing it more or less by accident, or are they only ‘the same thing under different descriptions’? Obviously a great deal of empirical research treats voluntary and involuntary behavior as causally equivalent and this practice has a lot of *prima facie* plausibility. As already noted, causal equivalence would be assumed if the efficacy of quinine were tested on persons who didn’t take it voluntarily. Similarly, an ‘internal’ study of the effect of quinine-ingestion on malaria attacks would be likely to treat the patients studied as ‘objects’, the actions of whose wills wouldn’t be considered in the study.

But it seems to me that the foregoing considerations are not entirely decisive. The question of whether there are real causal differences between doing things deliberately and merely experiencing them is in some ways parallel to questions that arose in classical physics about whether forces are real things. Forces won out in the struggle for a place within physical theory not because some particularly acute experimentalists succeeded in observing them, but rather because Newton succeeded in formulating precise and elegant laws involving them that enabled scientists to account for an amazing range of phenomena with amazing accuracy. Now, it seems to me that the question of the existence and nature of the will is similar.

I think that Berkeley’s argument against Newton, that *will* or spirit should take the place of Newton’s forces, is worth noting because Berkeley held that they have the same epistemological status as forces, namely as things that are not perceivable to the senses.\(^7\) This suggests that for the scientist, if not for the philosopher or the linguist, whether the will is to be accorded a place in theory ought to be decided not by observation or by semantic analysis, but by whether precise theories that accurately account for observational phenomena can be formulated in terms of it. It seems to me that current theories of choice can be interpreted as struggling in that direction, while on the contrary, of course, a lot of recent theorizing in cognitive science struggles against it.

Now we will briefly run through the exercise of formalizing the main elements of the theory that was set forth in the previous sections. This will also bring out connections with probabilistic theories of decision, concerning which we will make a few unsystematic comments.
11. Basic Concepts

We will focus on the quinine and malaria case as before, hoping that its transfer to other cases like smoking and cancer will be obvious. In this case there are two possible choices open to a person \( p \) on an occasion \( o \), to take the quinine or not to take it, which we will symbolize as \( q_1 \) and \( q_2 \), respectively, and two ‘outcomes’ to be considered, to suffer a malaria attack or not to suffer one, which we will symbolize as \( a_1 \) and \( a_2 \), respectively.

Note that \( q_1 \) and \( q_2 \) are possible acts while \( a_1 \) and \( a_2 \) are possible experiences, but our formalism doesn’t ‘show’ this difference.

Note too that, voluntary or involuntary, both the \( q_i \) and the \( a_j \) can be assumed to have intrinsic utilitarian values for \( p \), since other things being equal \( p \) would prefer \( a_2 \) to \( a_1 \) and \( q_2 \) to \( q_1 \); i.e., \( p \) would prefer not to have a malaria attack but also, probably, not to take the quinine. But some things lack intrinsic utility.

We will assume that there are two causal states that have instrumental value, inasmuch as their presence or absence influences the chances that \( p \) will have a malaria attack. These are: \( c_1 = p \) has malaria, and \( c_2 = p \) does not have it. That we say that these have only instrumental and not intrinsic utility means that we are assuming that what cause us direct pain or pleasure are symptoms like malaria attacks, and not their underlying causes, which are only valued indirectly, in terms of their effects. The chance of an ‘event’ \( \phi \) will be symbolized \( C_{p,o}(\phi) \), where \( \phi \) is any logical combination of \( a_j \) and \( c_k \), and subscripts \( p \) and \( o \) are appended because these quantities may depend both on the person, \( p \), and on the occasion, \( o \), when the chances apply.

Since \( C_{p,o}(\phi) \) will be assumed to be a probability function we also consider the conditional chance, \( C_{p,o}(\phi/\psi) \), that \( \phi \) will occur, given \( \psi \). \( C_{p,o}(a_1/q_1) \) corresponds to the proportion of ‘malaria attack’ balls in the ‘take quinine’ urn in front of individual \( p \) on occasion \( o \), and that this depends on \( p \) and \( o \) corresponds to the fact that these proportions may change from person to person and from occasion to occasion. Now we formulate two laws of chances.
12. Chance-Independence and Chance-Inequality Conditions

It is plausible that the states $c_k$ are causally independent of the $q_i$, i.e., whether or not $p$ has malaria is independent of her taking the quinine. This is formalized as a Chance-Independence Condition:

\[ \text{IndC} \quad C_{p,o}(c_k/q_1) = C_{p,o}(c_k/q_2). \]

On the other hand, the result of taking quinine is dependent both on the person’s causal state $c_k$ and on her action, $a_j$, and this can be stated as:

\[ \text{PCG} \quad \begin{align*}
C_{p,o}(a_1/q_1 & c_1) < C_{p,o}(a_1/q_2 & c_1), \quad \text{and} \\
C_{p,o}(a_1/q_1 & c_2) &= C_{p,o}(a_1/q_2 & c_2) = 0.
\end{align*} \]

for any person $p$, occasion $o$, and causal state $c_k$. That is, if $p$ has malaria on occasion $o$ then she will be less likely to have an attack of it if she takes quinine than if she doesn’t, and if she doesn’t have malaria her chances of having an attack are 0, whether or not she takes the quinine. The important thing is that she can choose to take the quinine and thereby influence her chances of having a malaria attack, though PCG makes it clear that the chances depend on other factors as well, in particular on whether she actually has malaria. Of course PCG holds at most in limited generality, and the practical business of science is to evaluate it. But how it should be evaluated necessarily depends on its use, which Section 7 suggested has to do with the long run and with laws of large numbers.

13. The Long Run, and a Law of Large Numbers

Now consider sequences of person-occasion pairs,

\[ \Gamma = \{ (p(1), o(1)), \ldots, (p(n), o(n)) \}. \]

For $m = 1, \ldots, n$ let $C_{p(m), o(m)}(a_1)$ and $C_{p(m), o(m)}(a_2)$ be, respectively, $p(m)$’s chance of having a malaria attack or not having one on occasion $o(m)$, and let $f_r(a_1)$ and $f_r(a_2)$ be, respectively, the frequencies with which persons $p(1), \ldots, p(n)$ experience or don’t experience such attacks on these occasions. Also, let

\[ \text{MC}_\Gamma(a_j) = \sum_{m=1}^{n} C_{p(m), o(m)}(a_j), \]
be the *mean conditional chance* of persons \(p(m)\) suffering malaria attacks \((j = 1)\) or not suffering them \((j = 2)\) on occasions \(o(1), \ldots, o(n)\). Then laws of large numbers entail, roughly, that, provided chances are independent, for \(j = 1, 2\) and any positive \(\varepsilon\), the probability that

\[
\text{LLN}_\Gamma(a_j) \quad |f_r(a_j) - MC_\Gamma(a_j)| < \varepsilon
\]

approaches 1 as \(n\) approaches infinity (this will be generalized in the following section). Note that this bears out the point that the mean chances \(MC_\Gamma(a_j)\) are not empirical frequencies, and certainly not the very frequencies, \(f_\Gamma(a_j)\), that also enter into \(\text{LLN}_\Gamma(a_j)\), for if they were then \(\text{LLN}_\Gamma(a_j)\) would be an uninformative triviality. Nor are they subjective quantities, since if persons’ degrees of belief were systematically at variance with the objective chances, their mean values might not converge to the same limit, and \(\text{LLN}_\Gamma(a_j)\) would be false for them. For \(\text{LLN}_\Gamma(a_j)\) to make sense, \(MC_\Gamma(a_j)\) must be a mean of objective chances that persons can influence, for instance by taking quinine.

The foregoing fits in well with the urn model. As already suggested, the conditional chances \(C_{p(m), o(m)}(a_j), j = 1, 2\), correspond to the proportions of balls marked ‘malaria attack’ and ‘no malaria attack’, respectively, in the urns that person \(p(m)\) chooses on occasions \(o(m)\), for \(m = 1, \ldots, n\). The frequencies, \(f_\Gamma(a_j), j = 1, 2\), that enter into \(\text{LLN}_\Gamma(a_j)\) correspond to the proportions of ‘malaria attack’ and ‘no attack’ balls that persons \(p(m)\) draw on occasions \(o(m), m = 1, \ldots, n\), and \(\text{LLN}_\Gamma(a_j)\) itself says, roughly, that as \(n\) grows large the proportion of sequences \(\Gamma\) in which \(f_\Gamma(a_j)\) differs by less than \(\varepsilon\) from the mean of the proportions of ‘malaria attack’ or ‘no attack’ balls in the urns actually chosen must approach 1.

Now we can consider the long-run gains that a person may aim to maximize by acting on prudently calculated chances.

14. *Expected Long-Run Gains and Their Maximization*

Now consider personal sequences \(\Gamma = \{p, o(1)\}, \ldots, \{p, o(n)\}\), involving a single person, \(p\), where the chances \(C_{p, o(m)}(a_1)\) depend on \(p\’s\) actions on occasions \(o(1), \ldots, o(n)\). If \(p\) takes quinine on all of these occasions then the chances \(C_{p, o(m)}(a_1)\) will all equal \(C_{p, o(m)}(a_1/q_1)\), while if she doesn’t take it on these occasions then \(C_{p, o(m)}(a_1)\) will all
equal \( C_{p,o(m)}(a_1/q_2) \). Then, letting

\[
MC_{\Gamma}(a_1/q_k) = \sum_{m=1}^{n} C_{p(m),o(m)}(a_1/q_k),
\]

we can generalize law \( LLN_{\Gamma}(a_1) \) to say that, assuming independence of chances, the probability that

\[
LLN_{p,\Gamma}(a_1/q_k) \quad |f_{\Gamma}(a_1/q_k) - MC_{\Gamma}(a_1/q_k)| < \varepsilon
\]

must approach 1 as \( n \) approaches infinity, where the frequencies \( f_{\Gamma}(a_1/q_1) \) and \( f_{\Gamma}(a_1/q_2) \) are, respectively, those of malaria attacks that \( p \) would suffer if she took quinine on occasions \( o(1), \ldots, o(n) \), and if she didn’t take it on those occasions. Note that when we consider frequencies \( f_{\Gamma}(a_1/q_1) \) and \( f_{\Gamma}(a_1/q_2) \), and the application of the law of large numbers to both of them, we are moving from its application to \textit{actual} series to \textit{hypothetical} ones, since clearly both series cannot be actual.

In either case, the frequency \( f_{\Gamma}(a_1) \) of malaria attacks that \( p \) can expect to suffer on occasions \( o(1), \ldots, o(n) \) will almost certainly be near to the mean chances \( MC_{\Gamma}(a_1/q_k) \), which gives \( p \) a practical influence on these chances. Moreover it follows immediately from \textbf{PCG} that \( MC_{p,\Gamma}(a_1/q_1) \leq MC_{p,\Gamma}(a_1/q_2) \), hence \( p \) will minimize her mean chances of having a malaria attack if she chooses \( q_1 \), to take the quinine on occasions \( o(1), \ldots, o(n) \), i.e., always taking the quinine will almost certainly minimize the frequency with which \( p \) has malaria attacks in the long run, so that if she is solely interested in this, she should always take it. In this way \textbf{PCG} can guide practical action, not necessarily to \( p \)’s immediate benefit, but rather to her long run advantage. This also shows the practical importance of \textbf{PCG}’s being \textit{right}, for if it should be mistaken then acting on this guidance would not be expected to yield \( p \) this long run benefit. Furthermore, this can be fine-tuned, as hinted at in Section 7.

When \( p \) has other interests and values the picture is complicated by having to quantify and balance them, and a simple way to factor in a cost, say of taking quinine, is as follows. Suppose that for individual \( p \), having a malaria attack on occasion \( o \) has value \( V_{p,o}(a_1) \) and not having one has value \( V_{p,o}(a_2) \), and taking quinine has value (or cost) \( V_{p,o}(q_1) \) and not taking it has value \( V_{p,o}(q_2) \), and suppose that these values remain constant over occasions \( o(1), \ldots, o(n) \).
Now it follows from \( \text{LLN}_{p,\Gamma}(a_1/q_k) \) that taking quinine on these occasions will almost certainly maximize the ‘mean payoffs’ for \( p \) if the following inequality is satisfied:

\[
(q_1 > q_2) \quad MC_{p,\Gamma}(a_1/q_2) - MC_{p,\Gamma}(a_1/q_1) > \frac{V_{p,o}(q_2) - V_{p,o}(q_1)}{V_{p,o}(a_2) - V_{p,o}(a_1)}.
\]

In words: if the difference between the mean chance of suffering a malaria attack if quinine is not taken and the chance of suffering one if it is taken exceeds the \textit{value-difference ratio}, i.e., the ratio of the difference between the ‘cost’ of taking the quinine and that of not taking it and the difference between the value of suffering a malaria attack and not suffering one, then \( p \) should take the quinine.\(^{15}\) If the cost of taking quinine is negligible in comparison to that of suffering a malaria attack the difference in the mean chances is likely to exceed it and \( p \) should take the quinine. On the other hand if \( p \) doesn’t have malaria and is certain not to suffer an attack of it, then the mean chances of having an attack are 0 whether or not she takes quinine; their difference is 0, and she shouldn’t take the quinine no matter how small the value-difference ratio is.

The general subject of policies for maximizing gains in the long run is obviously a very complex one, which won’t be pursued here.\(^{16}\) But it should be clear that rules like \( q_1 > q_2 \) involve \textit{objective} mean chances, independent of what persons may think the chances are, and they describe policies for action that would be best from the point of view of these people’s \textit{subjective} values. Many recent decision theories assume that similar precepts describe rational action in ignorance of chances, where subjective ‘degrees of belief’ take the place of objective chances and ‘one-shot gain’ takes the place of long-run advantage. I can only record here my skepticism as to whether these theories describe behavior that is practically advantageous, and therefore rational in any sense that it is desirable to be. That they may give rough, first-approximation descriptions of actual human behavior is to be explained, I think, by the fact that they describe the properties of objective chances, and persons roughly conform to them because objective chances conform to them and persons aim to bring their degrees of belief into conformity with these chances.\(^{17,18}\)

Finally, we will return briefly to estimation, i.e., to bringing our degrees of belief into conformity with objective chances.
15. Estimation Again

A decision maker, \( p \), using a practical causal generalization like ‘quinine wards off malaria attacks’ to guide her actions needs to estimate hypothetical mean chances such as \( MC_{p, \Gamma}(a_1/q_2) \) and \( MC_{p, \Gamma}(a_1/q_1) \) that enter into \( (q_1 > q_2) \). According to \( \text{LLN}_{p, \Gamma}(a_1/q_k) \), these will be, respectively, very probably near to the frequencies, \( f_{\Gamma}(a_1/q_1) \) and \( f_{\Gamma}(a_1/q_2) \), of malaria attacks that \( p \) would suffer on a series of occasions \( o(1), \ldots, o(n) \) if she took quinine and if she didn’t take it. However, all that applying laws of large numbers to data of past associations between quinine-taking and malaria attacks in sequences \( \langle p(1), o(1) \rangle, \ldots, \langle p(n), o(n) \rangle \) guarantees is that the frequency of malaria attacks suffered by \( p(1), \ldots, p(n) \) on occasions \( o(1), \ldots, o(n) \) are probably close to the mean chances that \emph{those} persons had for suffering them on those occasions. Only if it can be assumed that observed past series and the persons in them \emph{resemble} person \( p \) and her hypothetical future series can data from the former serve as reliable guides to the latter.

But laws of large numbers do not justify such inductive, ‘uniformity of nature’ assumptions. I believe that basing estimates on carefully selected series such as might be derived from test and control groups, and other techniques of experimental and statistical method, are inductive methods that go beyond simple laws of large numbers. In doing this they improve on our naive preconceptions about statistical uniformity, and about the connection between correlations and causality. But this paper doesn’t aim to improve on the methodology of estimating chances; it only aims to explain the advantage of being right about them, and their connection with a kind of generalization that guides so many of our actions.

ACKNOWLEDGEMENTS

I am indebted to Professors Brian Skyrms and Wolfgang Spohn for very helpful comments on an earlier draft of this paper.

\textit{Department of Philosophy,}
\textit{University of California at Berkeley,}
\textit{Berkeley, CA 94720,}
\textit{U.S.A.}
NOTES

1 This suggests the plausibility of a counterfactual formulation of non-practical causal generalizations: A’s cause or increase the chances of B’s if, and only if, the chances of B’s would be greater if A’s were brought about than if they weren’t. But this is only a suggestion, and it must be reiterated that our present concern is not with impractical generalizations – and also not with singular causal propositions, or with explanations.

2 This would be like discovering the ‘Springs of Nature’ that Hume spoke of. Of course this assumes that the chance of drawing a particular kind of ball from an urn is equal to the proportion of balls of that kind in the urn, which is not necessarily the case.

3 Strictly, if the aim is only to determine whether the person will be less likely to have a malaria attack if she takes quinine than if she doesn’t, which, on the urn analogy, would be equivalent to determining whether there is a smaller proportion of ‘malaria attack’ balls in the ‘take quinine’ urn than in the ‘no quinine’ urn from which the person must choose, it should only be necessary to compare the urns ‘back to back’. However, the statistical methods considered here require estimating proportions separately as a preliminary to comparing them, much as though the comparative heights of people could only be determined by measuring them independently, rather than by standing them back to back.

4 Essentially, that the test and control groups should be representative in the sense that the mean proportions of ‘malaria attack’ balls were the same in the ‘take quinine’ urns in both groups, and the same held for the mean proportions of these balls in the ‘no quinine’ urns in the two groups.

5 If any freedom of choice were left to members of the test group it could still happen that only persons with severe cases of malaria took quinine, and that would bias the results of the test.

As an aside it may be noted that dividing into test and control groups is a kind of partitioning, and that Suppes–Cartwright screening off, which is a kind of partitioning (cf. Skyrms, 1980, Ch. IIB), may derive its practical utility from the assurance that it provides that statistical results may be useful to many individuals. For instance, one might partition by age and sex in order to test the efficacy of quinine in warding off malaria attacks in individuals of each of these types. This aspect of statistical testing will not be considered in this paper.

6 We mustn’t think that shifting the inquiry to the object, the action, is shifting it to another practical causal relation, namely that between the will and the action. While this relation may be causal it isn’t practical, since persons don’t choose to will.

7 “I say, lastly, that I have a notion of spirit, though I have not, strictly speaking, an idea of it” (3rd Dialogue, my italics).

8 In strictness, we should say that $q_1$, $q_2$, $a_1$, and $a_2$ stand for propositions to the effect that certain things are done and that certain events occur. Then propositional combinations can be formed, like $q_1 \& a_2$, expressing the fact that person $p$ takes quinine and does not suffer a malaria attack on occasion $o$.

9 Richard Jeffrey’s theory (Jeffrey, 1983) which among current probabilistic decision theories is most closely related to the ideas developed here, follows this practice because it doesn’t recognize a distinction between actions and other kinds of happenings. The present theory could have differentiated formally between acts and ‘states’ as L. J. Savage’s theory (Savage, 1954) does, or by using a different font for ‘action
propositions’. This would have the advantage of showing clearly that we don’t assign probabilities or chances to propositions of those kinds.

10 In effect, this transfers the traditional distinction between what is ‘immediately perceived’ and what is only ‘inferred’ to values.

11 Strictly speaking, if \( q_i, a_j, \) and \( c_k \) are individual happenings then they depend on \( p \) and \( o \) as well, since, for instance, taking quinine on one occasion isn’t the same action as taking it on another. However, this fine point will be neglected here.

12 But a chance like \( C_{p,o}(a_j/q_i) \) need not be a ratio of unconditional chances, \( C_{p,o}(a_j&q_i)/C_{p,o}(q_i) \), since both \( C_{p,o}(q_i) \) and \( C_{p,o}(a_j&q_i) \) may be undefined; i.e. unconditional chances of action propositions may not be defined. Jeffrey (1985, Section 1.2) describes the use of ‘non-fractional conditional probabilities’ in a formalism for probabilistic choice.

13 We will use mnemonic labels like ‘\textbf{IndC}’ rather than simple numbering for displayed equations and related conditions.

We might worry that assuming that having malaria is independent of taking quinine would imply that taking quinine is independent of having malaria – that equating causal influences with inequalities of conditional chances would entail that causal influences are symmetric because chance-inequalities are. But this would be a mistake, since saying that quinine taking is positively correlated with malaria only means that among different persons, those who take quinine are more apt to have malaria than those who don’t, while saying that having malaria is independent of quinine taking means that an individual’s chances of having malaria are the same, whether or not he or she takes quinine. This obviates the need to introduce extraneous factors like time order into the analysis in order to ‘desymmetrize’ the causal order, though it by no means explains our grounds for thinking that we cannot influence the past. Note that in the case of practical causal generalizations, temporality is built directly into the causal object, the action.

14 This could be expressed informally as “taking quinine reduces a malaria sufferer’s chances of a malaria attack”, which is still a practical causal generalization.

15 It is easily shown that if the hypothetical mean chances in \( q_1 < q_2 \) are replaced by subjective conditional probabilities, say \( P_{p,o}(a_1/q_1) \) and \( P_{p,o}(a_1/q_2) \), the resulting condition follows say from Jeffrey’s decision theory on the assumption of ‘value additivity’:

\[
V_{p,o}(q_i&a_j&c_k) = V_{p,o}(q_i) + V_{p,o}(a_j),
\]

so long as it is assumed that the expected values of both the \( V_{p,o}(q_i) \) and the \( V_{p,o}(a_j) \) are zero, which is not implausible. This is itself a special case of value-independence,

\[
V_{p,o}(q_i&a_j&c_k) = V_{p,o}(q_i&a_j),
\]

which makes the assumption that the states \( c_k \) have no intrinsic utility explicit, and it is also compatible with so called ‘causal’ decision theories, to which Jeffrey’s theory actually reduces on the assumption that \textbf{IndC} holds. Conversely, Jeffrey’s theory can be interpreted in terms of objective chances, but then it would be misleading to describe the theory as ‘epistemic’, since there is no necessary connection between these chances and people’s states of knowledge or belief.

16 I haven’t yet even established whether \( q_1 < q_2 \), would still be valid if the constant values it involves were replaced by their corresponding mean values, in circumstances
in which the values $V_{p,o}(q_i)$ and $V_{p,o}(a_j)$, $i, j = 1, 2$, vary with the occasions $o(l)$, $o(n)$.

17 Something like this is argued for in Section 6 of Adams (1988), and this is closely related to remarks in Sections 4 and 5 of Ramsey (1926). We will not enter here into the thorny question of whether the 'subjective probabilities' or 'degrees of belief' of the subjectivists are best interpreted as estimates of chances.

18 If this is right it has implications for the controversy between so called 'epistemic' (R. C. Jeffrey, 1983, Eells, 1991) and 'causal' (e.g., Harper and Gibbard, 1978) schools of decision theory. Both schools are subjectivist, but if the suggestion made above is right their disagreement really comes down to the question of what laws are satisfied by objective chances, and of \textsc{IndC} in particular.

REFERENCES


COMMENTS BY PATRICK SUPPES

I have been discussing and arguing philosophical questions with Ernie Adams for at least four decades and for me the pleasure of these conversations has continued unabated. Reading his paper and commenting on it is like having another such conversation, and I am sure that my comments here will not be the last words we will have to say to each other about causality, probability, utility and related matters.
Much of what Ernie has to say I find congenial and sympathetic to my own approach to many causal questions. I particularly like what he has to say about the parallel between Newton’s analysis of forces and the way we can think about doing things deliberately. In other words, I am quite sympathetic with his general views about the primacy of experience of intentional actions. There are certain points, however, on which we still find ourselves apart, either on formulation or on a direct substantive issue.

**Objective Chances.** My main differences with Ernie’s approach concern his use of objective chances and the central place he assigns them. Towards the end of the paper he has very strong negative ideas on the use of subjective degrees of belief or subjective probability. I find this extremely puzzling, especially in the context of dealing with practical actions. Take the case of quinine and malaria that he discusses extensively. It could certainly be a matter of objective fact that many aspects of malaria are genetically related, not simply infection with the disease, but also the severity of attacks and the probability of self-immunization once infected. However, let us for the purposes of my analysis conjecture that any genetic components of this kind will not be identified for another hundred years. Furthermore, we can assume that the genetic factors produce considerable variation in individuals, so it may well not be possible for an individual who, for example, lives in an area heavily infested with malaria, to make anything like a correct decision according to objective chance. Watching other people around her suffering from malaria, this person may decide the only reasonable thing to do is to take quinine, even though there are some mild side-effects from doing so. In fact, the quinine is an absolute waste, for this person already has immunity from childhood malaria and has no chance of an attack, however she, and in earlier years her parents, do not want to take the risk, and so she continues to take quinine. Now the concept of being rational in my own view is ordinarily used very much in relation to the knowledge available to us, not to God, when we make decisions. We could of course use all kinds of frequency data about other people taking quinine, but we cannot approach the question of a detailed and serious biochemical theory of immunity, because the knowledge is not available to us at this time, even though we may suspect there is such a theory with practical consequences. I framed the question so that for an individual person, the deviation from objective chance in terms of the
estimation of chances or subjective probabilities that are used has real practical consequences. It seems to me that it is a criticism of Adams’ ideas that he does not seem to want to recognize the wide prevalence in these medical matters of such individual discrepancies between objective chances that could not possibly be known without taking risks that the individual does not want to take, or, even with taking the risk, can still not possibly be known.

Other kinds of examples, not involving cancer or malaria but involving actions, make me even more skeptical that in many cases we have a serious access to frequencies, let alone frequencies in the long run, but we still take what we regard as rational actions. In many parts of the world, the weather varies widely over short periods of time and with great unpredictability. Smith is faced with deciding whether to take his umbrella to work or not. He looks outside or even possibly listens to the radio to get the latest weather forecast. On the basis of that information and experience from the past, but not experience in any way encoded explicitly in terms of frequencies, he decides to take his umbrella. He may well have taken a rational action on the basis of the information available to him, but he had no access to frequencies, and certainly no access to objective chances. Yet it seems to me that in a great variety of cases of action, in fact I would be willing to claim in the majority of cases of action, this is exactly what we are faced with. We do not have objective chances available in any serious form. We must proceed with limited information, limited resources of all kinds, but still make a decision. Of course, good subjectivists use frequency data that they judge of good quality whenever possible to improve their subjective estimations of the probability of phenomena occurring.

The Transcendental. I am puzzled by Adams’ strong phenomenological conception of the transcendental. To paraphrase Berkeley, Adams seems to want to hold that what is out of sight is transcendental. Certainly causes or forces in general are not available to our senses, but some kinds of forces certainly are, in the form of pressure on our body parts, and similarly we all find ourselves accepting as knowable many things about other individuals that are inferred from their behavior, including their facial expressions, tone of voice, etc. To call all these matters transcendental and perhaps going so far as to say ‘out of sight, then transcendental’ is too limiting and too constrictive on our ordinary ways of thinking. I am sure Ernie does not really think of any one of
his children as being transcendental once they are a few hundred miles away, because they are no longer within his range of sense experience.

To take a really strict view of that which is given as phenomena, is to get us back into the old problems of sense data and other positivistic difficulties. Our ordinary language and ordinary thinking, as well as our scientific language and thinking, are saturated with concepts that depend upon inference from that which is immediately present to our senses. His remark at the end of Section 6 about the causes of malaria or cancer being transcendental is surprising, even though he wants to explain why causes are still important. The way he put the matter in the context of his extended discussion is that almost all the actions we think of as serious depend upon inferences to what he terms transcendental causes. The reason is that our actions are inextricably linked with our conceptual analysis of our own causal structure as well as the causal structure of the environment around us.

I turn now to some minor comments.

**Estimating Frequencies.** I liked Adams’ comment that use of the law of large numbers is probably not nearly so important as detailed experiments of various kinds. In this context, I have a skeptical query about his statement that we cannot estimate both the conditional frequencies of taking quinine and not taking quinine for a given individual. This is in Section 14. But it seems to me that this is exactly the kind of thing that he might want to consider, given his general views, namely that what will be run is what is called an inrasubject experiment with random periods of using quinine and not using quinine to see if there is a difference in the frequency of attacks under the two conditions for the same individual. In fact, detailed experimentation of this kind for single individuals would be the way to get at the genetic conjecture I made above in first approximation.

**Events as Causes.** I am puzzled about Ernie’s wanting so severely to limit causal analysis to actions. In taking actions we often need to make causal estimates of events causing other events. Because I believe that the rain caused the flooding in my basement and I believe that rain will occur again soon, I take certain actions that I would not otherwise. Such a setup seems to me an important aspect of a large number of the actions we take. We take actions because of our beliefs about the great variety of
events that will cause other events that we want to prevent or help occur.

**Practical Importance of Causality.** I very much agree with the main thrust of Section 3, that it is important to identify causes that take us beyond correlation because causes can be at work in a positive way, even when correlations are negative. This is the point that Laplace emphasized many years ago in insisting that causal analysis can show how causes change probability, not necessarily produce causes or produce results of certainty. It is a point sometimes not adequately appreciated. Being able by appropriate conditionalization to change a strong negative correlation to a conditional one that is weakly negative or positive can be of the greatest importance, as Ernie rightly stresses.