H. KYBURG

SUBJECTIVE PROBABILITY: CRITICISMS, REFLECTIONS, AND PROBLEMS

1. INTRODUCTION

The theory of subjective probability is certainly one of the most pervasively influential theories of anything to have arisen in many decades. It was developed first by probability theorists and philosophers (Koopman and Ramsey, primarily); then by a few somewhat unconventional statisticians (De Finetti and Savage). Growth in interest in the topic among statisticians was slow at first, but turned out to be (it seems) exponential in character. From statistics it spread to economics, political science, and the social sciences in general. From philosophy it spread to psychology and decision theory, and thence again to economics and political science. Although one could not say that it was dominant in any of these fields, it occupies a respectable place, either as subject matter or as methodological doctrine, in each of them.

One reason for this spread is both cultural and practical. The theory of subjective probability is undogmatic and anti-authoritarian: one man's opinion is as good as another's. It imposes only relatively weak constraints on the beliefs that a man has; if two people disagree, their disagreement can be treated merely as a datum, rather than as a cause of conflict. The theory embodies a modern spirit of tolerance and democracy. At the same time, in many areas people's degrees of belief are in relatively close agreement. The theory can also simply accept this agreement as a datum, and go on from there.[0] There is no need to justify that agreement, or even to look into its source. We need not get hung up on deep and abstract (and possibly artificial) issues in order to proceed to use this theory of probability in testing hypotheses, in making statistical inferences, in formulating theories of rational decision, in describing choice behavior under uncertainty, and so on.

The reason for the philosophical popularity of the doctrine is that it appears to be minimally committal – that is, it is easy to say that whatever the true doctrine of probability may be, *at least* it requires that the ordinary

Journal of Philosophical Logic 7 (1978) 157–180.

of the probability calculus hold of it; and since it is just these axioms that are required by the subjective theory, anything we can establish using the subjective theory will, *a fortiori*, be established for a more demanding theory. The theory provides a good working basis for philosophical investigation because it is maximally tolerant and minimally restrictive.

In the pages that follow, I shall argue that although the theory appears to be all things to all people, in fact it is a snare and a delusion and is either vacuous and without systematic usefulness, or is simply false. Before charging off into the brush, however, let us get straight about the animal we are after. By 'the' theory of subjective probability, I mean any of quite a large group of interpretations of the probability calculus which lead to the assignment of a numerical probability to each sentence or proposition of a language (or event of a field of events), and in which these assignments do not reflect any known or hypothetical frequencies.

The language in question may be restricted to the part of a natural language in which sentences relevant to a given problem or decision may be formulated, or it may be a larger or smaller fragment of an explicitly formalized language. Sometimes (particularly in statistical applications) the propositions or sentences are parametrized, and we consider, not the assignment of probabilities to specific propositions, but rather the assignment of probability densities to the parameters.

What do the numbers represent? This depends on the way in which the theory is construed. It may be construed as a descriptive theory or as a normative theory; and it may be construed as a theory of decision under uncertainty, or as a theory concerned with degrees of belief. The interpretations offered by various authors may not always fall cleanly into one of the resulting four classifications, but if the theory does not work under any of the four pure interpretations, it is unlikely that it can be saved by ambiguity.

If the theory is construed as a descriptive theory of decision-making under uncertainty, the numbers represent theoretical psychological characteristics (conveniently, if loosely, referred to as degrees of belief) which enter into a general descriptive theory of actual behavior under uncertainty. If the theory is construed as a descriptive theory of degrees of belief, the numbers represent actual degrees of belief, which we may attempt to measure by examining behavior under uncertainty. If the theory is construed as a normative theory of decision, the numbers represent parameters which *ought* to satisfy certain constraints. If the theory is construed as a normative theory of degrees of belief, then the numbers represent the measures assigned to sentences of a language, and *ought* to satisfy at least the constraints of the probability calculus. I include among the latter theories such 'logical' theories as that sought by the early Carnap, according to which on logical grounds alone we should be able to assign a probability measure to every sentence of the language, those discussed by the later Carnap, in which the probability measures of the sentences of the language are determined by a few parameters that represent personal or subjective commitment, and those proposed by Hintikka, Tuomela, Niiniluoto and others in which the probability measures of the sentences are determined by a few parameters which reflect *empirical* 'judgments' about the actual nature of the world.

All of these interpretations impute the same structure to probabilities (relevant exceptions will be noted in due course):

$p(h)$ is a real number lying between 0 and 1

$p(e \wedge h) = p(e) \cdot p(h/e)$, where $p(h/e)$ is the conditional probability of h, *given* e.

If h and g are logically exclusive, $p(h \vee g) = p(h) + p(g)$.

2. THE DUTCH BOOK ARGUMENT

In subjectivistic interpretations of probability, there is a fundamental nexus in which logic and behavior meet. It is the betting situation, first described with the intent of clarifying the relation between probability and belief by F.P. Ramsey [20]. Basically, the connection is this:

Step I: The more convinced I am that something is going to happen, the higher the odds I am willing to offer in a bet that it will occur. Ramsey's idea was to take the *least odds* at which an individual would accept a gamble on the occurrence of E to indicate his degree of belief in E. We assume that these least odds are well defined: that is, we assume that if the agent is willing to accept a gamble at $P{:}S$, then he will accept a gamble at $P'{:}S$ for any P' less than P. This can be made more precise and more plausible as follows:[1] Consider a gamble for a fixed stake S where a price P is charged. The least odds at which the agent is willing to accept such a gamble is the least upper bound of the ratios $P/(S-P)$ representing the gambles the agent is willing to accept. We also assume that these least odds are independent of the size of the stake. We do not assume, yet, that the least odds will remain unchanged for a negative stake – i.e., for bets on the non-occurrence of E.

We do assume, however, that the individual has least odds for a whole algebra of events (or propositions), i.e., that if his least odds are well defined for E and F, they are well defined for $\sim E$ and for $E \wedge F$.

Step II: We regard it as irrational for a person to have a book made against him – i.e., to accept a sequence of bets such that, no matter how events turn out, he loses.

Step III: We conclude that it is irrational for a person's beliefs to be represented by least odds such that the person can be the victim of a Dutch Book concocted by an unscrupulous bettor. (I don't know who first used the word 'unscrupulous' in this context, but it does create a persuasive image of the virtuous Rational Man protecting himself against the Forces of Evil.)[2] From this, in turn, it follows that those odds must be related in just the way that probabilities are related in the standard probability calculus. Whatever else is true about rational beliefs, then, (we might, for example, demand with Carnap that they reflect certain linguistic symmetries) they must at least satisfy the core axioms of the probability calculus. We have thus found, in the Dutch Book Argument, a justification for the demand that rational degrees of belief satisfy these axioms. (If we are construing the subjectivistic interpretation of probability empirically, we have found in the Dutch Book Argument *prima facie* warrant for taking the probability calculus as a plausible empirical hypothesis.)

Or have we? It is noteworthy that Carnap, in *The Logical Foundations of Probability* [1], does *not* employ these arguments to justify the axioms he lays down for degree of confirmation. He was perfectly aware of Ramsey's work, and devotes several pages to defending Ramsey against the charge of 'subjectivism', there construed as the denial that there is a difference between merely coherent degrees of belief and *rational* degrees of belief. Ramsey himself writes, "It is not enough to measure probability; in order to apportion correctly our belief to the probability, we must also be able to measure our belief." (Quoted by Carnap [1], p. 16.) Of course we know that Carnap was later persuaded that the Dutch Book Argument was sound, and on a subjectivistic view it may very well turn out to make no sense to attempt to measure probability and belief independently. Nevertheless, let us begin by taking a close look at the argument.

We first note that Step II has nothing to do with degrees of belief or with probabilities. No rational person, whatever his degrees of belief, would accept a sequence of bets under which he would be bound to lose no matter what happens. No rational person will in fact have a book made against him. If we consider a sequence of bets, then quite independently of the odds at which the person is willing to bet, he will decline any bet that converts the sequence into a Dutch Book. His least odds on E, for example, may be 2 : 1, while his least odds on E and F may be 4 : 1; this violates the calculus, but is no Dutch Book. His least odds on E may be 2 : 1, and his least odds on $\sim E$ may be 2 : 1; there is not even a Dutch Book here, unless he accepts each bet – and, of course, being rational, he will accept no more than one bet, for otherwise, at the same amount on each bet, he would be bound to lose.

According to the subjectivistic theory, it is irrational to offer 2 : 1 odds that a certain toss of a coin will land heads and to offer 1 : 1 odds that it will land tails. But it is *not* unreasonable, on this view, to offer 2 : 1 odds to A that the toss will land tails, and then to offer 2 : 1 odds to B that the following toss will land heads. Nor is it irrational to make the following bet: I agree to draw a card from an ordinary deck, and if it is red offer you 2 : 1 odds on heads, and if it is black, offer you 2 : 1 odds on tails. In fact, however, so far as the Dutch Book argument goes, it would be perfectly rational for me to offer you 2 : 1 odds on heads *and* 2 : 1 odds on tails, provided I make it clear that I am offering only one bet. Note that this has nothing to do with my placing a value on gambling: If I *accepted* both bets, it would imply that I valued the gamble in itself; but my refusal to accept more than one bet precludes warrant for the assertion that I am assigning a high value to the utility of gambling.

The rational man will not actually have a book made against him; but this is a matter of deductive logic, and not of probability; a matter of certainties and not a matter of degrees of belief. Step II then is a heuristic device, or, to put it less charitably, a red herring.

The conclusion that a rational man's degrees of belief, as reflected in his least odds, must satisfy the axioms of the probability calculus does not follow from the assumptions thus far made.

What we require in order to capture the intent of the Dutch Book Argument is an intermediate step:

Step I-A: We assume that the individual is willing to take any combination of bets, for any stakes, at the least odds that characterize his degrees of belief.

In effect, this step amounts to the demand that the individual post odds on a field of events or statements, and then *make book* according to the odds he has posted. Note how different this is from Step II: Step II points out that a rational man will not *actually* have a book made against him. But we already know that, and that has nothing to do with the 'least odds' at which he would bet on an event, or even on each event in a field of events. This consideration imposes no constraints on the numbers we assign to statements or events. Step I-A, on the other hand, imposes a constraint on the posted odds concerning a whole field of events, when the individual is required to accept all bets offered to him at those posted odds. The constraint here concerns *potential* sets of bets, not actual ones. The rational man, in posting his odds, protects himself against all possibilities; he must make book with all comers, as an equal opportunity bookie, according to the odds he has posted. Thus we must replace the assumption of Step II by the following assumption:

Step II-A: The rational man will not put himself in a position in which he can *potentially* have a book made against him.

Note, however, that the conclusion that these posted least odds can be converted into a set of numbers satisfying the axioms of the probability calculus still does not follow. The least odds the agent offers on E may be 2:5, and the least odds he offers on $\sim E$ may also be 2:5. In fact, if he is a professional gambler, the sum of the odds he offers on any two complementary bets will of course add up to less than 1; this is how he makes his money. And there is hardly anything irrational about this!

We may, however, arrive at the following conclusion: with every event or proposition we may associate two numbers, one representing the least odds at which the agent will bet *on* the event, the other representing the least odds at which the agent will bet *against* the event. We can now show that there exists a function P, satisfying the axioms of the probability calculus, such that for every event in our field of events the value of that function will belong to the closed interval determined by the corresponding numbers.

A number (a relatively small number) of writers follow the Dutch Book Argument only this far; C. A. B. Smith [25], I. J. Good [7], P. Suppes [26], I. Levi [15], A. Dempster [4] are among them. It is possible then to take the set of functions P satisfying the posted odds to characterize an individual's degrees of belief. Note that in so doing we are abandoning the notion that there is a single real-valued degree of belief associated with each proposition in our algebra, determined by the least odds at which an individual will bet *on* that proposition.

What we require in order to obtain the conclusion of the Dutch Book Argument is a stronger assumption by far than any we have made so far. It may be stated thus:

Step I-B: We assume that the individual is willing to take any combination of bets, for any stakes, *positive or negative*, at the least odds that characterize his degrees of belief.

To stipulate that the odds may be positive or negative has the effect of requiring that the agent be willing to take either side of a bet at his posted odds. Now, at last, the conclusion of Step III will follow from the assumptions of Steps I, I-B, and II-A. The posted odds can now be translated into numbers, associated with each member of the field in question, which satisfy the probability calculus. There are other ways to obtain such numbers which eliminate the need to be sloppy about the marginal utility of money and the desirability of gambling: for example we may force (or bribe) an individual to express a preference between acts that depend on the states of nature (the field of events). The best known approach along these lines is that of Savage [20].

What is noteworthy is that we obtain a set of numbers satisfying the probability calculus only by *compelling* the individual to do something: to post odds on E and odds on $\sim E$ and then to accept any bets, with positive or negative stakes, at these odds; or to express a preference between receiving a prize if it rains on March 3, 1986, and receiving a prize if between 1,456 and 1,603 heads appear on 3,341 tosses of a specific coin. It is only through this sort of compulsion that we can obtain a full set of numbers satisfying the probability calculus.

But now we see that the argument from Step I, Step I-B, and Step II-A, to Step III, while valid, is no longer sound: we have no reason to suppose that an individual would be willing to post odds under these conditions, or to take the time and effort to express a serious set of preferences. We must replace I-B by:

Step I-C: The individual can be *made* to post odds on which he will take any combination of bets, for any stakes, positive or negative.

It will follow that these odds will conform to the probability calculus. But now the connection between these odds and degrees of belief has become attenuated to the point of obscurity. However irrational and strange my degrees of belief, I will, under compulsion, post odds that are coherent (or publish a coherent preference ranking).

As I look out the window, it occurs to me that the least odds I would be comfortable about offering on rain today are about 3 : 7, and the least odds I would be comfortable about offering against rain are about 3 : 7. If I am forced, in accordance with Step I-C, to post odds and make book, I will pick a number between 0.3 and 0.7 to determine those odds. But I will be no more comfortable with 0.5 than with 0.4 or 0.6. The number I pick will depend on my mood and my circumstances. But it does not seem to me that my 'degree of belief' in rain is varying correspondingly according to my mood and circumstances.

There may indeed be a rough intuitive connection between my degree of belief in an event E and the least odds I am willing to offer in a bet on E. But this connection is much too loose to generate by itself a set of numbers conforming to the probability calculus. The Dutch Book Argument gives excellent reasons for adopting a table of odds or publishing a list of preferences which conform to the basic core of the probability calculus, but it does so at the cost of severing the immediate and intuitive connection between odds and degrees of belief that the argument originally depended on. In fact, at this point, we may find ourselves wondering if there *is* any such thing as 'degree of belief'.

There are a number of directions we can go from here, and various theories of subjective probability have explored these various directions. We can regard 'degree of belief' as a kind of intervening variable in an empirical decision theory: that is, a psychological theory that accounts for the decisions that people actually make. The theory would assert that people act as if they had degrees of belief conforming to the probability calculus, and so acted as to maximize their expected utilities. Or we can regard 'degree of belief' as a kind of intervening variable in a normative decision theory: people *ought* to act as if they had degrees of belief conforming to the probability calculus, and were maximizing their expected utility. Or we can suppose that decisions and preferences just constitute a way of *getting at* degrees of belief (or that there is some other way of getting at degrees of belief), and construe the probability calculus as a theory of people's actual beliefs. Or, finally, we can suppose that decisions and preferences are just ways of getting at degrees of belief (or that there is some other way of getting at degrees of belief) and construe the theory of subjective probability as a normative theory of degrees of belief. In the sections that follow, we shall explore each of these four alternatives in turn.

3. EMPIRICAL DECISION THEORY

A number of investigators have taken subjectivistic probability as an ingredient in an empirical theory of how people actually make decisions. People's decisions, of course, are highly complex, and involve a number of factors that are extremely hard to evaluate. The theory is tested, therefore, under relatively simple and artificial circumstances. For example, subjects are presented with a board containing a large number of red and white thumbtacks. Although the subject is not given enough time to count the thumbtacks, it is found that usually his judgment of the proportion of red thumbtacks is quite accurate. The subject is offered a pair of bets on red against white. He chooses a bet. A row and column of the thumbtack display are selected by a device which insures that each thumbtack will be selected equally often (and the subject knows this). The bet is then settled. The bet is sometimes a matter of real money; sometimes a matter of hypothetical real money bets; sometimes a matter of chips or counters [5].

There are a number of things to be noted about such experiments. First, subjective factors that might be regarded as extraneous to the probability–utility framework are excluded as rigorously as possible. Second, the 'amounts' involved in the bets are arranged so that they are quite strictly linear in utility. Third, although the probabilities involved are 'subjective', they very closely reflect the frequencies of the various outcomes. Thus these experiments constitute a simple and direct test of the SEU (subjective expected utility) theory of decision under circumstances highly favorable to the theory. As Edwards remarks, if the theory is going to work for subjective probabilities in general, it should certainly work for subjective probabilities which simply reflect objective frequencies [5].

It turns out that while the theory predicts reasonably well for probabilities in the middle range, it goes awry for large or small probabilities, despite the fact that the utilities involved are well within the linear range. Furthermore, it goes awry in different ways for positive expected utilities and for negative expected utilities.

Even in these highly artificial and simple situations, people do not act in ways that are consonant with the assumption that their degrees of belief, as reflected in their choices between bets, satisfy the axioms of the probability calculus.

From a psychological point of view, one is interested in developing a theory of choice behavior under uncertainty which will enable one to predict and understand people's choices, not in devising a theory of belief that embodies certain *a priori* constraints. Psychologists therefore seem to have largely moved away from the SEU model. Much of the recent literature relevant to this development is cited in [24], especially pp. 9–11.

That psychologists are abandoning the SEU model does not mean in itself that subjective probability is being abandoned. One possibility is that 'subjective probability' – i.e., degree of belief – may no longer be required to satisfy the axioms of the probability Calculus. Edwards [5] suggests the possibility that the additive property be abandoned. The abandonment of any such property entails the rejection of the Dutch Book Argument as yielding a description of choice behavior even under ideal circumstances. It strongly suggests that it is inappropriate to call the 'subjective probabilities' that enter into such a revised empirical theory 'probabilities' at all.

Another possibility that is being actively pursued is that an accurate descriptive decision theory will involve not only expected gain, but other moments of the gain as well. Thus people might be concerned not only with their *expectation* in a gamble, but with the *variance* of the outcome, prefering, say, gambles with modest expectation and small variance to gambles with larger expectations but much larger variance. Here again, however, we find problems for the conventional subjectivistic interpretation of probability. The conventional interpretation supposes (as in the Dutch Book Argument) that we can get at subjective probabilities in a straightforward way through compulsorily posted betting odds, or through the analysis of forced choices among alternatives. But this approach takes the SEU model for granted, and applies it inside out, so to speak, to determine subjective degrees of belief. If we suppose that preferences are not determined by subjective expected utility, but in some other more complicated way, it may be difficult to measure 'degrees of belief'.

In any event, the neat relations among 'degree of belief', 'utility', and 'expected utility' that underly the Dutch Book Argument are not reflected in people's actual choice behavior.

4. NORMATIVE DECISION THEORY

One response to this situation (I think it would have been Savage's, and it would certainly be that of a large number of philosophers and economists) is to say that the subjective expected utility model of decision was never intended to be an empirical description of how people actually make decisions, even under 'ideal' circumstances, but rather a normative prescription of how people ought, rationally, to make decisions. Of course we suppose that people by and large are relatively rational – and the evidence of the psychological experiments for the midrange of probabilities provides gratifying confirmation of the general rationality of people. But the fact that when the probabilities become (according to the theory) large or small people cease somewhat to be rational should not be construed as *undermining* the normative theory.

Let us, then, confront this normative theory in its own terms, ignoring the fact that people are not always rational in their actual behavior. Indeed, were people always rational in their decisions, we would hardly be motivated (except perhaps as metaphysicians) to develop an elaborate theory of rationality. In point of fact, the main use of normative decision theory is in providing guidance under circumstances of uncertainty. Savage himself remarks that ". . . the main use I would make [of the axioms] is normative . . ." ([23], p. 20).

There are a number of ways of formalizing normative decision theory in a subjectivistic framework. The following, adapted from Savage [23], will suffice for our purposes. We need to get at both utilities and probabilities; the mechanism for doing so is to look at preference rankings. Preference rankings of what? Of *acts*, construed as functions from states of affairs to consequences. We can choose our acts – that is what decision theory is all about – but we cannot choose the state of the world, nor can we choose the consequences of our acts. Normative decision theory is designed to guide our choices among acts.

The subjective approach supposes that on the basis of an individual's ranking of acts (functions, remember) we may infer both his probabilities and his utilities. It must be supposed that probabilities and utilities cannot be evaluated independently of the ranking of acts – else the agent might be in the position of ranking most highly an act whose mathematical expectation is not the highest, and this blatantly violates the basic thrust of the subjectivistic approach. On the other hand, if we begin with a complete preference ranking among acts, what is the point of the analysis into utilities and probabilities? The process seems obviously circular: we start with a preference ranking among acts, and by dint of careful analysis arrive at probabilities and utilities, and by computing mathematical expectations arrive at a ranking which (unless something has gone wrong) precisely matches the original one!

The point, however, is that the individual's intitial preferences may be partly unknown and partly incoherent. For example, he may prefer act A to act B; act B to act C; and act C to act A. This is clearly irrational: preference should be transitive, and nothing should be preferred to itself. His preference ranking may be incoherent in a more sophisticated way: for example, he may prefer A if state X obtains and B if state not-X obtains to A if X, and he may prefer A if X to B if not-X. This set of preferences is incoherent because there is no assignment of probabilities and desirabilities that will lead to mathematical expectations that conform to this ranking. (Suppose that the probability of X is p; then the value of the first ranked alternative is p times the value of the second ranked alternative plus $(1-p)$ times the value of the third ranked alternative; the value of the mixed alternative must be between the values of the pure alternatives, since $0 < p < 1$; this contradicts our original stipulation.)

If such an incoherence is pointed out to a rational individual, he will presumably alter his preference ranking. Savage says that to be in this state is to be 'uncomfortable' in the same sense as to find oneself committed to a logical inconstency. But to be 'in this state' is to presuppose that one's preferences form a simple order, and as Savage himself recognizes, actual preferences form at best a partial ordering [23, p. 21]. It is precisely in the attempt to make one's preferences conform to the theory – that is, in the attempt to produce a simple order of preferences from one's actual partial order – that one stumbles on violations of the theory. Unlike logical consistency, coherence hangs on a doctrine that is admittedly unrealistic; without the doctrine we might not find anything uncomfortable about failures of the transitivity of indifference, for example.

The normative force of the theory, then, is that it says that the preference ranking of an individual *should* be such that there exists a probability function and a utility function for which the mathematical expectations computed in accordance with those functions agree with the preference ranking.

Axiomatizations of subjective theories generally require an assumption even stronger than that of simple order among preferences. Savage [23, p. 6, p. 39] requires that there are states of the world with arbitrarily small probabilities. He is able to show that a preference ranking satisfying his postulates can be decomposed into a unique probability function and a unique desirability function. The desirability of an act can then be represented as its mathematical expectation.

Jeffrey [11] imposes a condition that is somewhat weaker on preference rankings: the 'splitting condition' requires that any (non-neutral) element of the preference ranking can be expressed as the disjunction of two equiprobable incompatible elements that are ranked together [11, p. 104]. Jeffrey's system does not entail that there are *unique* probability and desirability functions that expresses the preference ranking; it merely entails that there is a family of pairs of probability and desirability functions which will *fit* the preference ranking, in the sense that mathematical expectations computed in accord with any of them will yield rankings among elements that conform to the original rankings.

Now how do we use this normative decision theory? We begin with a partial preference ranking. The axioms of the theory help us to make our preference ranking more complete. That this sort of exercise can be very useful is clear; subjectivist decision theory has become entrenched in business schools precisely because of this usefulness. It is most useful when there are relatively standard measures of desirabilities (dollars) or of probabilities (frequencies) or both. There are often a fairly large number of fixed points in a preference ranking that can be established by such considerations; these can give some indication of how 'inconsistencies' are to be resolved. But the axioms cannot do the whole job; if we start with a partial ranking, there may well be acts whose relative positions in the preference ranking are not determined by the initial partial ranking and the axioms of the theory. We must then consult our intuitions to decide how those alternatives are to be ranked.

It may be the case that our initial partial preference ranking is incoherent – even though it is only partial, it may conflict with the axioms of the normative theory. As Savage points out on numerous occasions, the theory cannot tell us how to resolve that conflict. Again we must consult our intuitions. The theory tells us that some change must be made in our preference ranking, but it does not tell us what change. It will generally be the case that our preference ranking is incomplete in the sense that, given our initial preferences, there will be alternatives whose location is not determined by those preferences together with the axioms of the theory. Again, we must call on intuition to complete the preference ranking, to the extent that it needs to be completed. (In order to use the theory we need not always have a preference ranking complete in all detail.)

Once we have a preference ranking that satisfies the axioms of the theory, we can crank out at least one probability and desirability function, and perhaps a family of them. We can now use these probability and desirability functions to compute mathematical expectations. And we are then directed to perform that act with the greatest mathematical expectation. But of course that act is the act at the head of this coherent preference ranking. So if we've got a coherent preference ranking, we don't have to make any computations. On the other hand, if we don't have a coherent preference ranking, the theory won't tell us *how* to make it coherent. If we have a coherent but partial preference ranking, the theory may allow us to make it more complete. The theory thus functions as a heuristic device enabling us to specify our preference ranking in an organized way, and revealing incoherencies that require resolution.

We must now ask, however, how important the role of subjective probability is in enabling normative decision theory to perform its function. Our suspicions may be aroused by two facts: First, Jeffrey's form of the theory allows the derivation from the preference ranking of not one, but of a whole family of probability functions. If these functions are construed as yielding subjective probabilities – i.e., degrees of belief – this means that the preference ranking does not either yield or require a *unique* degree-of-belief function. Second, the most clear-cut and persuasive applications of the theory are those in which the subjective probability function simply mirrors well established statistical frequencies. In these applications, then, we can construe the function whose combination with a utility function is to yield the coherent preference ranking as a perfectly straightforward frequency function, having no psychological import whatever. (It *may* be the case that there are such things as degrees of belief, and it *may* be that in certain situations these degrees of belief have the same magnitudes as certain relative frequencies, but in the applications we are discussing both of these assumptions are gratuitous and irrelevant.)

Many of the most interesting of the applications of the normative decision theory, however, are applications which do *not* involve well tested roulette wheels or other apparatus yielding alternatives with well-known frequencies. The question now is whether or not the probability numbers that emerge from the analysis of the preference ranking can nevertheless be construed as known (or reasonably believed) frequencies. Certainly in many of the instances in which the normative decision theory is applied – in business decisions, for example – it is not hard to imagine that the probability numbers can plausibly be construed as estimates of the relative frequency with which a given sort of thing happens in given circumstances. Again, to the extent that this is plausible, there is no need to construe these numbers as reflecting subjective degrees of belief.

It may be questioned whether or not this is always the case. There are certainly applications of Normative Bayesian Decision Theory in which the probability numbers emerge from a dialogue between the decision theorist and the executive decision maker, rather than from any recorded tables of statistical data. But this does not settle the question of what the numbers that thus emerge represent. On the subjectivistic interpretation of probability, they represent degrees of belief; but we may also suppose that these numbers represent the executive's (possibly "subjective") estimates of *objective* frequencies. We might be hard put, on occasion, to formalize these intuitive judgments: it can be far from obvious what frequency in what reference class is being intuitively estimated when the decision maker says that he would just bet even money that the sales of product X in 1979 will exceed 145,500 units. Nevertheless, it is not implausible to suppose that there is some estimate of an objective frequency underlying such judgments.

It is worth remarking that the specification of a reference class is not a trivial problem even in the case of probabilities that are quite clearly related to frequencies: probabilities that arise in weather forecasting, in insurance, and the like. For formal languages there is a mechanism that will yield a correct reference class, given a body of knowledge – see [13]. The extent to which this mechanism can be valuable heuristically in relatively informal contexts remains to be seen.

To sum up: If we suppose we begin with a full preference ranking among acts, there are two possibilities. Either the preference ranking is coherent, or it is not. If it is coherent, we are all set – we merely follow the dictates of our preference ranking with no further analysis. If it is not, then something must be changed; but as Savage never tired of pointing out, the subjectivistic theory will not tell you what to change. Subjectivistically interpreted, Normative Bayesian Decision Theory, whatever its heuristic virtues, is either philosophically vacuous or impotent. What has made the theory attractive to non-philosophers – to psychologists and economists and businessmen – is of course precisely its heuristic virtues. But as I have tried to point out, *these* virtues are preserved – and quite possible enhanced – by adopting a point of view in which probabilities are given an interpretation which rests ultimately of frequencies rather than on degrees of belief. Finally, since Jeffrey can achieve a representation of coherent preferences through the use of a whole family of pairs of utility and probability functions, it seems clear that it is unnecessary and perhaps ill advised to hypostatize these probabilities as 'degrees of belief' in even an idealized psychological sense.

5. EMPIRICAL THEORY OF DEGREES OF BELIEF

Can we construe the subjectivistic theory of probability as an empirical theory of degrees of belief? We note that the failure of the theory as an empirical theory of decision does not entail its failure as a theory of degrees of belief. The subjectivistic theory construed as a theory of decision making involves two other ingredients besides degrees of belief: it involves utilities, and it involves the acceptance by the individual of the principle of maximizing mathematical expectation. Thus the kinds of experiments that psychologists have performed to test the subjectivistic theory of decision are not really decisive with regard to a subjectivistic theory of degrees of belief. But then in order to test the latter theory we must find some way – preferably more direct – to measure degrees of belief than that to which we are led by the full blown subjective expected utility theory.

One way would be to inquire of people what their degrees of belief in various propositions are. This is a very unlikely approach: people will often say they don't know, and they will often announce numbers that do not fit into the calculus of probability. We must assume – if the theory is to have a chance at all – that people have no very clear access to their degrees of belief.

Furthermore, it is not altogether clear that there are degrees of belief. We can introspectively distinguish a number of qualities of our beliefs: confidence, enthusiasm, intensity; but it is not so easy to sort our beliefs out along the linear array that we want to think of as representing degrees of belief. It is particularly difficult to do this introspectively.

Savage has suggested a number of ingenious ways to get at degrees of belief which do not involve the entire subjectivistic theory of decision. For example, one could discover the maximum amount that a person would pay for a ticket that would return him a dollar if S is true; that amount would represent his degree of belief in S. But this will not do for several reasons. First of all, the prices a person offers for tickets on a number of related propositions may well not satisfy the axioms of the probability calculus. In fact this will quite generally be the case, since there are traditional probability calculations (the birthday problem, for example) that most people find surprising. Second, on the subjectivistic view, a person may change his opinions at any time, for good reason or for no reason. To discover whether or not a person's degrees of belief satisfied the axioms of the probability calculus would require that we test all of his beliefs – or a large set of them – simultaneously. But of course we can't actually do this. Finally, it is difficult to see how in *this* program of measurement we can eliminate the influence of risk: a conservative sort would, I imagine, discount his tickets (say, offering $p - \epsilon$ for a dollar ticket on S, and $1 - p - \epsilon$ for a dollar ticket on $\sim S$), and an enthusiastic gambler might well offer a little extra for the fun of gambling. The same sorts of problems arise here as arise in the general attempt to assess probabilities by gambling behavior discussed in Section 2.

Another device proposed by Savage, and endorsed by a number of writers on subjective probability, is the forced choice: the subject is asked to choose between receiving a substantial prize if it rains in Los Angeles three weeks from today, and receiving the same prize if a sequence of fifteen coin tosses contains 12 heads. Although no finite number of such forced choices can show that his degrees of belief *do* obey the probability calculus, a finite number of them can show that they don't. If the forced choices yield an ordering of propositions that can be explained by a set of degrees of belief satisfying the calculus, that is certainly evidence that the theory has something to it.

But we have every reason to expect that such a test would fail to support the theory. Just as the degrees of belief that people claim will not satisfy the probability calculus (else they would not be surprised by probability calculations), so their choices, which no doubt come close to reflecting the degrees of belief they would claim, will not be consistent with the probability calculus. And we have the same temporal problem as before: since on the subjective view, degrees of belief may change freely, we must measure a number of related degrees of belief simultaneously, and the proposed technique does not allow us to do that.

Finally, we may suppose that someone invents an epistemeter which will measure a thousand degrees of belief simultaneously through electrodes wired to the head of an individual. If there were such a device, it could put our theory to relatively direct test. The device would have to be tuned and standardized, of course, but that could perhaps be done by means of some of the techniques mentioned earlier. Now if we had such a machine, it is conceivable that then we could discover something much like degrees of belief that characterizes people's brains, and which does indeed satisfy the axioms of the probability calculus. But I wouldn't bet on it. At any rate, so far as *present* evidence is concerned, we have no evidence that people's degrees of belief satisfy the probability calculus, and considerable evidence to the effect that they do not.

6. NORMATIVE THEORY OF DEGREES OF BELIEF

Relatively few people, I assume, have seriously proposed that the subjective theory of probability be construed as an empirical theory of degrees of actual belief. They have said, rather, that the subjective theory is ideal or normative: should an individual discover that his degrees of belief do not conform to the probability calculus, he will just by that very fact be motivated to change them in such a way that they do conform. When someone shows me the calculation which establishes that the odds are better than even that that two people in a room of twenty-one people will have the same birthday, I revise my degree of belief accordingly. According to the introspective testimony of some, the feeling is much like that of being caught in a logical contradiction.

It is difficult to make much out of this introspective feeling; not all individuals testify to having felt it, and it hardly seems a firm enough foundation to build a normative theory of degrees of belief on. Ramsey [20] pointed out long ago that there was no point in a normative theory of belief if there were to be no way of measuring beliefs.

If we had an epistemeter, of course, we could apply the normative theory: we could take an individual, wire him up, and see whether or not his momentary degrees of belief conformed to the probability calculus; if they did, we could give him a gold star. Science fiction aside, however, there seems to be no behavioral way to get at degrees of belief. Even in the most persuasive case, when we demand that an individual post odds representing his degrees of belief, and take all bets offered, we have no assurance whatever that his degrees of belief will conform to the probability calculus, precisely because of the fact that if he is *deductively* rational, he will not post incoherent odds, however incoherent his degrees of belief may be.

At the moment, then, there is no way of getting at a person's degrees of belief which is dependable enough that it could serve as a ground for asserting that he is or is not rational in his degrees of belief. But we may nevertheless consider the subjectivistic theory as a standard of rationality that we are simply not in a position to *apply* practically at the moment. It may still be important theoretically and philosophically, and it may in fact turn out to be feasible to apply it practically some day. Indeed, this is the way in which most philosophers who take the subjectivistic theory of probability as a standard of rationality, more or less without argument, seem to look on it. It is simply assumed that whatever else may be said about rationality, at any rate one's degrees of belief should conform to the probability calculus.

There are several things to be said about this. The first and most obvious is that it is not at all clear that there *are* degrees of belief in the sense required by the demand that they conform to the probability calculus. The calculus requires that with each proposition there be associated a real number, and that these real numbers be related in certain ways. But there may be nothing that corresponds to the term 'degree of belief'; or if there is it may need to be measured by an interval, or by an n-dimensional vector. It seems reasonable to suppose that 'degree of belief' is a psychological concept which depends for its usefulness on the psychological theory in which it appears. But at the moment, there is no such psychological theory.

Second, this construal of the subjective theory entails that one should be very certain indeed of propositions that seem both very powerful and empirical. Thus if one has a degree of belief equal to p in a certain relatively rare kind of event (e.g., that the next ball drawn from an urn will be purple), and if one supposes that one's degree of belief that the second ball is purple, given that the first one is purple is no more than $2p$, and if one regards the draws as exchangeable in the subjectivist's sense, then one should be 99% sure that p plus or minus 0.01 of the draws, in the long run, will yield purple balls. (For details of such arguments, see [12].)

This is, indeed, what one's initial beliefs commit one to, if they conform to the probability calculus. On the other hand, most of us would regard the consequence as unintuitive. The moral, I believe, is that there is no way to make one's beliefs conform to the probability calculus – on the assumption that they are real numbers – without doing violence to some rational intuitions. To reject the consequence requires – essentially – rejecting the supposition that there are any exchangeable sequences of events, which would undermine the usefulness of the subjectivistic theory.

Finally, we may ask what functions this normative theory of degrees of belief performs. It is clear that it serves no function for non-philosophers. But is it possible that it performs a function for philosophers – can we get some mileage out of a theory which supposes that there are degrees of belief, and demands that they conform to the usual probability axioms?

We consider two cases. First, suppose that these are all the constraints there are on rational belief. Then any distribution of degrees of belief over the propositions of a language is as good as – as rational as – any other, provided the axioms of the probability calculus are satisfied. But the axioms are compatible with any degree of belief in any individual (non-logical) statement. No degree of belief in any non-logical proposition can, by itself, be ruled irrational. The set of degrees of belief in a number of related propositions are constrained, on this view, by the requirements of rationality. But what does this constraint tell us? It does not tell us that the rational individual will not in fact have a book made against him, for if we grant him deductive rationality we already know that he won't actually allow a book to be made against him, regardless of what his degrees of belief are.

It might be maintained, and would be by anyone who regarded the theory of subjective probability as providing insights into scientific inference, that its main function is dynamic: it is the changes in the probability function that are wrought by empirical evidence, through the mediation of Baye's theorem (or a generalization thereof) that give the theory its philosophical importance. The most frequently cited examples are the convergence theorems: Given that two people have degrees of belief that satisfy the probability calculus, and given that their degrees of belief satisfy some relatively mild constraints in addition to the coherence constraints, then, as empirical evidence accumulates, their distributions of beliefs will become more and more nearly the same. Of course there is nothing *in the theory* that requires that even the mild constraints will be satisfied for a rational person. But the really serious problem is that there is nothing in the theory that says that a person should *change* his beliefs in response to evidence in accordance with Bayes' theorem. On the contrary, the whole thrust of the subjectivistic theory is to claim that the history of the individual's beliefs is irrelevant to their rationality: all that counts at a given time is that they conform to the requirements of coherence. It is certainly not required that the person got to the state that he is in by applying Bayes' theorem to the coherent degrees of belief he had in some previous state. No more, then, is it required that a rational individual pass from his present coherent state to a new coherent state by conditionalization. Just as he may have got to his original coherent state by intuition, whimsey, imagination, evidence processed through Bayes' theorem, or any combination thereof, so he may with perfect rationality pass from his present coherent state to a future coherent state through any of these mechanisms. If he depends on Bayes' theorem, it is a matter of predilection, not of rationality. For all the subjectivistic theory has to say, he may with equal justification pass from one coherent state to another by free association, reading tea-leaves, or consulting his parrot.

This leads us to the second and more interesting case, in which the subjectivistic theory of degrees of belief embodies more than the constraint of coherence. The most obvious addition will be a principle that directs us to change our beliefs in accordance with Bayes' theorem: When we pass from one coherent epistemic state to another, as a result of acquiring evidence e, the probability function that describes our new state should be the conditional-on-e probability function of the old state. This is the principle of epistemic conditionalization, and is accepted, at least implicitly, by almost all writers who employ the subjectivistic theory.

The principle of conditionalization puts the theory in a new light. Thus supplemented, the theory makes the rationality of my beliefs at a future time depend not only on their coherence, but on their history. My beliefs at that time will be rational only if they are derived from my present beliefs by conditionalizing on the evidence that becomes available to me between now and then.

But if the rationality of my beliefs tomorrow depends on my having rational beliefs today, and conditionalizing to reach tomorrow's beliefs, it follows by the Relativity of Time that the rationality of my beliefs today depends on my have had rational beliefs yesterday, and having conditionalized on them to reach today's beliefs, or else that I can start being rational at any time – in particular, tomorrow, so that tomorrow's beliefs need not, after all, be based on applying conditionalization to today's beliefs. In the latter case we are back to pure subjectivism. But in the former case we are in Carnap's old position – the one he occupied before being seduced by the siren song of subjectivism: in theory and in principle a rational being would have beliefs that are precisely determined by his total body of evidental knowledge as accumulated from time zero to the present, and by the language he uses. Probabilities – degrees of belief – are logically determinate, and quite independent of what degrees of belief any actual person happens to have. The problem, then, as Carnap saw, is to determine the values of the absolutely prior probabilities – the probabilities that should be assigned to the statements of the language prior to any experience at all. We should assign these probabilities on the basis of *rational intuition*. It turned out that this was extremely difficult to do, and even Carnap, in his later years, began to doubt its possibility. The demands such a program imposes on rational intuition are simply too great.

A middle position is defended by Hintikka and some members of his school [9, 17, 18]. We define a logical probability measure on the sentences of a language, and assert conditionally that *if* this measure represents our prior belief, and *if* the principle of epistemic conditionalization is accepted, *then* such and such interesting results follow. But no attempt is made either to defend the prior measure assignment as demanded by rational intuition or to defend it as corresponding to anybody's actual beliefs. It is sometimes said that (if it did represent someone's beliefs) it would represent a 'presupposition' about the nature of the universe. Since no attempt is made to defend these presuppositions as rational, we would seem to be back in the realm of purely subjective theory; and since it is not argued that the measures in fact represent anyone's beliefs, we seem to be in a realm both hypothetical and subjective, from which we can learn little of either philosophical or practical import.[3]

7.

Despite the fact that subjectivistic probability is highly fashionable both in statistics and in philosophy, it appears to have serious shortcomings. We may account for its popularity in statistics by its heuristic role in teasing out the commitments that are implicit in an agent's preference ranking; but its role there is purely heuristic, and we need not assume that there is anything psychological that corresponds to the 'degrees of belief' that emerge from certain of those analyses. And note that in some analyses – Jeffrey's for example – a unique probability function doesn't even emerge as an auxilliary notion from the analysis. We may account for the popularity of subjective probability in philosophy in part by fashion, in part by laziness (it is easy to manipulate), and in part by the fact that it has few viable competitors. Nevertheless it is poor philosophy to adopt a false theory to achieve a certain end just because one doesn't know of a true theory that will achieve that end.

I conclude that the theory of subjective probability is psychologically false, decision-theoretically vacuous, and philosophically bankrupt: its account is overdrawn.

University of Rochester

NOTES

[0] In some cases this agreement can be explained by convergence theorems.

[1] This formulation is due to Levi, "On Indeterminate Probabilities" p. 413–414.

[2] It is interesting to note, as pointed out to me by Teddy Seidenfeld, that the Dutch Book against the irrational agent can only be constructed by an irrational (whether unscrupulous or not) opponent. Suppose that the Agent offers odds of 2 : 1 on heads and odds of 2 : 1 on tails on the toss of a coin. If the opponent is rational, according to the theory under examination, there will be a number p that represents his degree of belief in the occurrence of heads. If p is less than a half, the opponent will maximize his expectation by staking his entire stake on tails in accordance with the first odds posted by the Agent. But then the Agent need not lose. Similarly, if p is greater than a half. But if p is exactly a half, then the rational opponent should be indifferent between dividing his stake (to make the Dutch Book) and putting his entire stake on one outcome: the expectation in any case will be the same.

[3] It is worth noting, however, that Levi's proposals in [14] and [15] escape these criticisms, and most of those that follow. He supposes that the epistemic state of the agent is represented by a convex set of coherent probability functions, thus escaping criticism based on the alleged exactness of so-called "degrees of belief". More important, as a good pragmatist he does not demand that the rationality of an epistemic state depend on its history; he is concerned mainly with the rationality of changes from one epistemic state to another, and while he takes some of these changes to depend on conditionalization, he also admits changes that do not depend on conditionalization.

BIBLIOGRAPHY

[1] Carnap, Rudolf, 1950, *Logical Foundations of Probability*, The University of Chicago Press, Chicago.
[2] Carnap, Rudolf, 1968, 'Inductive Logic and Inductive Intuition', *The Problem of Inductive Logic*, Lakatos (ed.), pp. 258–267.
[3] Carnap, Rudolf, 1961, 'Inductive Logic and Rational Decisions', *Studies in Inductive Logic and Probability I*, Carnap and Jeffrey (eds.), pp. 5–31.

[4] Dempster, A., 1961, 'Upper and Lower Probabilities induced by a Multivalued Mapping', *Annals of Mathematical Statistics* **38**, 325–339.
[5] Edwards, Ward, 1960, 'Measurement of Utility and Subjective Probability', in Gulliksen & Messick, pp. 109–128.
[6] Edwards, Ward, 1962, 'Subjective Probabilities Inferred from Decisions', *Psychological Review* **69**, 109–135.
[7] Good, I. J., 1962, 'Subjective Probability as a Measure of a Non-Measurable Set', in Nagel, Suppes and Tarski, *Logic, Methodology and Philosophy of Science*, Stanford University Press, pp. 319–329.
[8] Harper, William L., 1975, 'Rational Belief Change, Popper Functions, and Counterfactuals', *Synthese* **30**, 221–262.
[9] Hintikka, Jaakko, 1971, 'Unknown Probabilities, Bayesianism and De Finetti's Representation Theorem', *Boston Studies in the Philosophy of Science* VIII, Buck and Cohen (eds.), D. Reidel, Dordrecht.
[10] Holstein, C. Stael von, 1973, 'The Concept of Probability in Psychological Experiments', *Logic, Methodology and Philosophy of Science IV*, Suppes *et al.* (eds.), pp. 451–466.
[11] Jeffrey, Richard C., 1965, *The Logic of Decision*, McGraw-Hill Book Co., New York.
[12] Kyburg, H. E., 1968, 'Bets and Beliefs', *American Philosophical Quarterly* **5**, 54–63.
[13] Kyburg, H. E., 1974, *The Logical Foundations of Statistical Inference*, Reidel, Dordrecht and Boston.
[14] Levi, Isaac, 1977, 'Direct Inference', *Journal of Philosophy* **74**, 5–29.
[15] Levi, Isaac, 1974, 'On Indeterminate Probabilities', *Journal of Philosophy* **71**, 391–418.
[16] Niiniluoto, Ilkka, 1972–73, 'Inductive Systematization: Definition and a Critical Survey', *Synthese* **25**, 25–81.
[17] Niiniluoto, Ilkka, and Raimo Tuomela, 1973, *Theoretical Concepts and Hypothetical-Inductive Inference*, Reidel, Dordrecht.
[18] Pietarinen, Juhani, 1972, *Lawlikeness, Analogy, and Inductive Logic*, North Holland, Amsterdam.
[19] Popper, Karl R., 1955–56, 'Two Autonomous Axiom Systems for the Calculus of Probabilities', *British Journal for the Philosophy of Science* **6**, 51–57.
[20] Ramsey, Frank P., 'Probability and Partial Belief', in Ramsey, *The Foundations of Mathematics*, pp. 256–257.
[21] Ramsey, Frank P., 1950, *The Foundations of Mathematics*, Braithwaite (ed.), Routledge & Kegan Paul Ltd., London.
[22] Salmon, Wesley C., 1970, 'Partial Entailment as a Basis for Inductive Logic', Essays in Honor of Carl G. Hempel Rescher (ed.), pp. 47–82.
[23] Savage, Leonard J., 1954, *Foundations of Statistics*, John Wiley, New York.
[24] Slovic, Paul, Fischhoff, Baruch, and Lichtenstein, Sarah, 1977, 'Behavioral Decision Theory', *Annual Review of Psychology* **28**, 1–39.
[25] Smith, Cedric A. B., 1961, 'Consistency in Statistical Inference and Decision', *The Journal of the Royal Statistical Society*, Series B (Methodological), Vol. 23, No. 1, pp. 1–37.
[26] Suppes, Patrick, 1974, 'The Measurement of Belief', *Journal of the Royal Statistical Society*, Series B (Methodological), Vol. 36, pp. 160–191.