THE JOURNAL OF PHILOSOPHY
VOLUME LXXVIII, NO. 6, JUNE 1981

A BAYESIAN THEORY OF RATIONAL ACCEPTANCE*

UNDER what conditions should a rational person accept a proposition? If you had to pick a question that has preoccupied epistemologists and philosophers of science, you could hardly do better than pick this one. Yet, from a Bayesian perspective, it is by no means obvious why it is a question anyone should care to ask. It is not that the Bayesians are philosophical philistines. No less than others, Bayesians appreciate the drive to provide systematic insight into the rational pursuit of human endeavors. What Bayesians find difficult to see is exactly which human endeavors theorists of rational acceptance are concerned about.

I

Some people have written as if the theory of rational acceptance were concerned, at least in part, with rational decision-making. They have suggested that a rational person *acts on* the propositions she accepts: if a rational person accepts, with respect to a decision problem, the claim that of all her options A will have the best outcome if chosen, then she will choose A. On this view, then, the theory of rational acceptance is designed, at least in part, to say when a rational person should be willing to act upon a proposition.[1]

* For their encouragement and criticism I would like to thank Burton Dreben, Allan Gibbard, Warren Ingber, Thomas Ricketts, Lawrence Sklar, and Joan Weiner. A fellowship from the Canada Council supported part of the research of which this paper was a product. A shorter, programmatic, version of this paper was read at the Eastern APA meetings in December, 1980.

[1] See, for example, R. B. Braithwaite, "The Nature of Believing," *Proceedings of the Aristotelian Society*, XXXIII (1932/33): 129-146, and "Belief and Action," *Aristotelian Society* suppl. vol. XX (1946): 1-19; C. W. Churchman, "Science and Decision Making," *Philosophy of Science*, XXIII, 3 (July 1956): 248-249; R. Rudner, "The Scientist *qua* Scientist Makes Value Judgments," *ibid.*, XX, 1 (January 1953): 3. It should be noted that some theorists of rational acceptance have explicitly disavowed this view. See, for example, R. M. Chisholm, *Perceiving* (Ithaca, N.Y.: Cornell, 1957), pp. 10/1; I. Levi, *Gambling with Truth* (New York: Knopf, 1967), pp. 7-11. Both offer arguments much like the one below.

0022-362X/81/7806/0305$02.60

Were acceptance a state of certainty, the view that rational persons act on the propositions they accept would be an unobjectionable fragment of a theory of rational decision. But, given that virtually no one seems to hold that accepting P entails being certain that P,[2] the view is defective.

Suppose a patient comes into a clinic suffering great pain. There are two drugs the doctor considers prescribing. If she prescribes drug α, most but not all of the patient's pain will be alleviated. If she prescribes drug β, one of two things will happen: either all the patient's pain will be alleviated or, if he suffers from an extremely rare allergy, he will die in minutes from an irreversible allergic reaction to drug β. Suppose that the doctor, on the basis of good—but not conclusive—evidence accepts the claim that the patient does not suffer from the allergy. And so, she infers rationally, prescribing drug β will produce the best outcome. Of course, she is not certain that prescribing the drug will not kill the patient. Nonetheless, the view above requires her, on pain of irrationality, to prescribe β anyway—deliberately to risk killing the patient just in order marginally to reduce the patient's pain.

The Bayesians have a diagnosis of what has gone wrong: what a person accepts cannot provide sufficiently subtle doxastic input to inform that person's decision problems. In fact, the Bayesian theory of rational decision dispenses entirely with references to what a person accepts: a person is rational to choose to perform an act A just in case there is no alternative act that bears greater *expected utility* than A. Roughly, the expected utility borne by an act is equal to the sum of the weighted desirabilities of the possible outcomes of the act—the desirability of each outcome being weighted by the degree to which the person is confident that the act, if performed, will have that outcome. Hence, on the Bayesian theory, grave risks are taken into account, whereas, on the view rehearsed above, the acceptance of a proposition may suffice to cast caution to the winds.[3]

[2] Although Isaac Levi now seems to espouse this view. In *The Enterprise of Knowledge* (Cambridge: MIT. Press, 1980), Levi develops a theory of what he calls "acceptance as evidence" or "knowledge"—where this is construed as a state of certainty. In a review of Levi's book (forthcoming in *The Philosophical Review*) I argue that there are good reasons for steering clear of the certainty view of acceptance. But, for the purposes of this paper, I will simply assume that the reader is already convinced.

[3] This, of course, is not in itself a vindication of the Bayesian theory of rational decision. There are, however, a number of writers who have provided deep and, in my opinion, compelling rationales for the theory. The locus classicus is L. J. Savage, *The Foundations of Statistics*, 2nd ed. (New York: Dover, 1972).

Of course, to say that it has no role to play in the theory of rational decision is hardly to damn the theory of rational acceptance. Rational decision-making has, after all, been at most a peripheral concern of theorists of rational acceptance. Their major interest has been to provide a systematic theory of the rational adoption of doxastic attitudes.

This is an interest close to the Bayesian heart. Indeed, one of the best-known results of the Bayesian theory of rational decision is a contribution to a systematic theory of rational doxastic attitude-adoption. [Let '1' represent the maximum degree of confidence one can have in the truth of a proposition, '0' the minimum. And, accordingly, let every proposition receive some number n, $0 \leq n \leq 1$, representing X's degree of confidence in the truth of that proposition, where X is an arbitrary agent. Call this assignment of numbers to propositions X's *degree-of-confidence function.* It is a consequence of the Bayesian theory of rational decision that, if X is rational (and endowed with sufficient logical acumen and computational ability), X's degree-of-confidence function will satisfy the axioms of the probability calculus.] Accordingly Bayesians are disposed to be sympathetic with the acceptance theorists' desire to contribute to the theory of rational doxastic attitude-adoption—provided, of course, that acceptance is construed as a state of confidence. The trouble is that there is no reasonable way to meet this proviso.

We have already noted that acceptance is not plausibly to be construed as a state of certainty. Accordingly, if it is a state of confidence, acceptance must be a state of confidence above some threshold—most plausibly, some threshold greater than or equal to .5. Putting this more formally,

(1) There is a number n, $.5 \leq n < 1$, such that, for any person X and proposition P, X accepts P if and only if X has a degree of confidence that P greater than n.

Now, most writing on rational acceptance has presupposed that, when it comes to accepting propositions, a rational person is subject to certain consistency constraints. Insofar as her logical acumen does not fail her, she must not accept an inconsistent set of propositions. Faced with a reductio of a set of propositions she accepts, she must abandon at least one proposition in the set. It is upon these minimal constraints that most philosophers have sought to build the more elaborate set of constraints that would constitute a theory of rational acceptance. In other words, most philosophers have assumed that, except where her logical acumen fails her,

(2) If X is rational, then
 (a) X accepts the conjunction of any propositions she accepts;
 (b) X accepts all the consequences of every proposition she accepts; and
 (c) X does not accept any contradiction.

But (1) and (2) together do not provide a sound foundation for the systematic study of rational acceptance. Rather, they immediately yield a pair of puzzles, each a reductio exploiting (1) and (2).[4] Let us look at the first of them.

Suppose we pick a value for n in (1): let n be equal to .9. Suppose that X is rational, and X is certain [and thus by (1), accepts the claim] that there is a one-thousand-ticket fair lottery in which one and only one ticket will win. Suppose that, accordingly, X has, for each ticket t_i in the lottery, a degree of confidence equal to .999 that t_i will lose. By (1) and our choice of n, she accepts with respect to each i, $0 < i \leq 1000$, the proposition that t_i will lose. Since X is rational, by (2a) she accepts the proposition that t_1 and t_2 and . . . and t_{1000} will lose and that there are one thousand tickets in the lottery of which one and only one will win. Let us suppose that it is as obvious to X as it is to us that this proposition entails a contradiction. So, by (2b), X accepts a contradiction. But then, by (2c), X is not rational—which contradicts our supposition that she *is* rational.

The argument will obviously work for any value of n less than .9. And if we modify the case to increase the number of tickets in the lottery, the argument will work for any value of n greater than .9 but less than 1.

Now the second puzzle.

Suppose X is rational and a professional historian. X has, after long work and research, written a piece of history which now fills a rather hefty book. As she reads over her magnum opus she reaffirms, for each sentence in the book, her acceptance of the proposi-

[4] The first reductio is a variant of one due to H. Kyburg, Jr., *Probability and the Logic of Rational Belief* (Middletown, Conn.: Wesleyan UP, 1961), p. 463, and C. G. Hempel, "Deductive-Nomological vs. Statistical Explanation," in H. Feigl and G. Maxwell, eds., *Minnesota Studies in the Philosophy of Science*, vol. 3 (Minneapolis: Univ. of Minnesota Press, 1962), pp. 163-166. The second reductio is a variant on the one named by the title of D. C. Makinson's "The Paradox of the Preface," *Analysis*, xxv, 6 (June 1965): 205-207.

tion that that sentence expresses in English. But, having finished reading, she recognizes that she has written a great many things and that, given the ambition of her work and her human fallibility, her book is very likely to be erroneous in some detail or other. So *X* has a very low degree of confidence—below .5—that the conjunction of the propositions expressed in her book is true. Since *X* is rational and since, for each proposition in her book, *X* accepts that proposition, it follows by (2a) that *X* accepts the conjunction of all the propositions expressed in her book. But, by (1), since her degree of confidence in the truth of that conjunction is below .5, *X* does not accept that conjunction. And this contradicts what has just been claimed.

Insofar as we feel compelled by the two reductios, on pain of irrationality, either to abandon (1) or to abandon (2), we must recognize that we cannot do the latter. For it is only because we are committed to (2) that reductios *do* compel us to revise our assumptions. After all, in constructing a reductio, we do nothing more than derive a contradiction from the conjunction of a set of propositions we accepted (or entertained accepting) at the outset. Were we not committed to accepting the conjunction and consequences of what we accept—and accepting no contradictions—such a derivation could hardly have the critical force it has. Indeed, those who would conclude from our two reductios that we should reject (2) are guilty of a double mistake.[5] Not only are they licensing rational persons to ignore the force of reductios in general; they are licensing *us* to ignore the force of the very reductios they are using to convince us to reject (2).

Our only choice then is to abandon (1). That is to say, we must conclude that acceptance, being neither a state of certainty nor a state of confidence above a threshold, is not a state of confidence at all. But, from a Bayesian perspective, we thereby admit that acceptance is not a doxastic attitude. Thus, from a Bayesian perspective, a theory of rational acceptance cannot be offering a contribution to the systematic study of the rational adoption of doxastic attitudes—for acceptance is not a doxastic attitude.

So Bayesians are puzzled: if neither the theory of rational decision nor the theory of rational doxastic-attitude adoption is the proper concern of the theory of rational acceptance, what *is* its concern? In the absence of any answer from the community of epis-

[5] And a number of philosophers, most notable of whom is Henry Kyburg, have concluded just that. See for example, Kyburg's "Conjunctivitis," in M. Swain, ed., *Induction, Acceptance and Rational Belief* (Boston: Reidel, 1970), pp. 55-82.

temologists and philosophers of science, many Bayesians have concluded that talk about rational acceptance is irrevocably loose talk which is undeserving of serious systematic treatment; that serious epistemological inquiry should focus on the investigation of rational degree-of-confidence functions.

It may seem easy to dismiss the Bayesian critique. After all, it looks as though the Bayesians have just illicitly stipulated that the term 'doxastic attitude' applies exclusively to degrees of confidence and then have concluded that, given the indefinability of acceptance talk in terms of degree-of-confidence talk, acceptance is not a doxastic attitude. But this misses the real point of the argument.

Bayesians have worked hard to show why their favored way of talking about doxastic attitudes deserves an important place in a theory of rational persons. They have done so by showing how a rational person's degree-of-confidence function, together with her utility function, determines her decisions. Virtually no theorist of rational acceptance has even tried, let alone succeeded, to do likewise for *her* favored way of talking about doxastic attitudes. Acceptance theorists have yet to provide a credible story about how a rational person's acceptance of a proposition makes any difference at all to the way in which she will conduct her life—whether within the context of inquiry or without. The point of the Bayesian argument is that, given this fact, and given the indefinability of acceptance talk in terms of confidence talk, we have no account of why acceptance talk deserves any place at all in a theory of rational persons.

The Bayesian critique, in short, issues a challenge to theorists of rational acceptance: *either* to say what human endeavor they are worried about, say what difference a rational person's acceptance of a proposition will make to the way in which she engages in that endeavor, and show how to provide a systematic treatment of the rational pursuit of that endeavor, *or* to give up the theory of rational acceptance and do epistemology the Bayesian way.[6]

It is a challenge that I think theorists of rational acceptance should take seriously. It is also a challenge that can be met—or, at least, so I will argue. My aim in the balance of this paper will be to say how we should construe, and how we can systematically inves-

[6] The most eloquent exponent of this sort of critique has been Richard Jeffrey. See his "Valuation and the Acceptance of Scientific Hypotheses," *Philosophy of Science*, XXIII, 3 (July 1956): 237-246; "Probable Knowledge," in I. Lakatos, ed. *The Problem of Inductive Logic* (Amsterdam: North-Holland, 1968), pp. 166-180; and "Dracula Meets Wolfman: Acceptance vs. Partial Belief," in Swain, *op. cit.*, pp. 157-185.

tigate, the subject matter of the theory of rational acceptance in the light of the Bayesian critique. My aim, thus, is to offer a Bayesian theory of rational acceptance.

II

My contention is that the proper purpose of acceptance talk is quite different from the proper purpose of degree-of-confidence talk. We need degree-of-confidence talk to describe the doxastic input into rational decision-making. The proper function of acceptance talk, on the other hand, is to describe a certain feature of our behavioral repertoire—a practice in which I am engaging in the very writing of this paper.

What I am doing in writing this paper is defending a set of propositions. Insofar as I make myself clear, you, as reader, will infer that I accept these propositions. Notice that this inference is reasonable only because of the context in which you are reading this paper. Were this not a paper in a scholarly journal, but rather a paper in a collection of satire, you would make very different inferences. You might in that case suspect that I accept very little of what I am defending—that I am defending what I do with tongue in cheek. If you do not harbor such suspicions as you read this paper, it is because, given where the paper is appearing, you assume that my primary aim in writing what I do is not to amuse you or entertain you or shock you, as it might be were this an attempt at satire. You infer that I accept the propositions I am defending here because you assume that my primary aim in defending what I do is to defend the truth.

You are also prepared to make inferences in the other direction. Insofar as you suspect, for some proposition *P*, that I accept *P* even though I have not explicitly defended *P* here, you will also suspect that, were I asked whether *P* is true and were my primary aim to defend the truth, I would defend the claim that *P* is true. In short, you are prepared to say that I accept a proposition *P* just in case you are prepared to say that I would defend *P* were my aim to defend the truth. My suggestion is this: that we view '*X* accepts *P*' as nothing more than shorthand for '*X* would defend *P* were her aim to defend the truth'.[7]

[7] It is worth noting that I am using the verb, 'to defend', in only one of its senses. In the sense I am using, to say that *X* defended *P* is *not* to say that *X* offered a defense of *P*—say, by offering an argument for *P*. Rather, it is to say simply that *X* asserted that *P* or assented to *P*. My attempt to make good Bayesian sense of acceptance talk was inspired by R. B. DeSousa's "How to Give a Piece of Your Mind: or, The Logic of Belief and Assent," *Review of Metaphysics*, xxv, 1 (September 1971): 52-79, especially pp. 62/3.

The truth is just the true and comprehensive theory of the world, where a theory is *comprehensive* just in case, for every proposition *P*, the theory either entails *P* or entails the denial of *P*. The aim to defend the truth, taken quite literally, is not then an aim that anyone has achieved or expects to achieve. By "the aim to defend the truth" I mean, rather, the aim to defend as comprehensive a part of the truth as one can. As one would imagine, in pursuing this aim, the desire for truth and the desire for comprehensiveness often conflict. When one is choosing what to include in one's theory of the world, the choice between including a stronger rather than a weaker claim often requires *either* giving rein to the desire for comprehensiveness at the expense of increasing the likelihood that one's theory will be false *or* overriding that desire so as not to increase that likelihood. Thus a theory of rational acceptance, in my sense of 'acceptance', is a theory that says how a rational person ought to adjudicate between these oft-competing desires so as to decide what she should be disposed to defend when her aim is to defend the truth—i.e., what she should accept.

Although, on this view of acceptance, acceptance is not a doxastic attitude in the Bayesian sense, it is nonetheless tolerably clear what '*X* accepts *P*' says about the way in which *X*, if rational, is disposed to behave: it says that *X* would defend *P* if *X*'s aim were to defend the truth. It is also clear that a theory of rational acceptance, in my sense of 'acceptance', has a bona fide claim to our attention. After all, the practice of defending propositions as if one's aim were to defend the truth is, at least to us as investigators, a very important part of our behavioral repertoire. We devote a great deal of time and energy to the wealth of books, lectures, and teaching which are the conspicuous products of this practice. And we manifest great interest in what investigators are willing to defend and in the propriety of their willingness to defend what they do. But it is far from obvious that the theory of rational acceptance, so conceived, is susceptible to systematic treatment.

Once we endorse (2), our two puzzles impose important constraints on an adequate theory of rational acceptance. If *X*'s degree-of-confidence function in the lottery case is indeed rational (as it seems to be), then, by (2), there will be at least one proposition predicting an outcome of the lottery which *X* is required not to accept *despite* being rationally all but certain that it is true. And if (as also seems true) *X* in the history-book case is within the bounds of rationality to accept each of the propositions expressed in her book, then she is required by (2) to accept their conjunction *despite* being rationally all but certain that this conjunction is false.

I suspect that few will welcome these constraints. Indeed, to most, the prescriptions noted above will seem entirely unpalatable. And herein lies the trouble. If we can neither stomach the consequences of endorsing (2) nor stomach the consequences of rejecting (2), we seem to have little choice but to conclude that we should give up hope of constructing a systematic theory of rational acceptance at all.

Fortunately, reflection on what we have accomplished so far should suffice to show us why this is a conclusion we are not forced to draw. For our definition of 'acceptance' not only saves the theory of rational acceptance from Bayesian attacks upon its legitimacy; it also saves the theory from itself. Notice that the unpalatability of the prescriptions that issue from (2) derives almost entirely from our habit of thinking of acceptance as a state of confidence. On the view of acceptance being proposed here, however, the first prescription does *not* prohibit *X* from being extremely confident, with respect to each ticket, that this ticket will lose. Nor does the second prescription require *X* to become confident that everything she said in her book is true. Rather, the prescriptions say only that, when her aim is to defend the truth, *X* should *not,* with respect to each ticket, defend the claim that this ticket will lose, and *X should* defend the claim that everything she said in her book is true. Elsewhere,[8] I have shown how we can sketch a story about why, when adjudicating between the competing aims of truth and comprehensiveness, a rational person would choose to follow these prescriptions.

The problem of concern, however, is whether we can give anything more than a sketch. We certainly do not want to construct a theory that will prohibit defending all but the propositions of which we are certain. Nor do we want to construct a theory that will endorse defending every garden-variety improbability. We will have to construct a theory that will say how a rational person adjudicating between the desire for truth and the desire for comprehensiveness could distinguish between defensible probable propositions and indefensible ones and between defensible improbable propositions and indefensible ones.

The project of the next section is to present just such a theory. But before we begin, let me make a few preliminary remarks.

On the view of acceptance we are taking here, there is a division of labor among epistemological theories: the theory of rational degree-of-confidence functions determines the rationality of one's

[8] In "Rational Acceptance," *Philosophical Studies* (forthcoming).

doxastic input into decision-making in general, and the theory of rational acceptance determines, given a doxastic input, the rationality of defending a proposition when one's aim is to defend the truth. So, for the purpose of pursuing a theory of rational acceptance, we can and will assume that the preliminary concerns peculiar to the theory of rational degree-of-confidence functions have been met.

Our vehicle in that pursuit will be an idealized agent X. X has complete logical acumen and computational ability, and X is rational in the Bayesian sense: X has a degree-of-confidence function that satisfies the axioms of the probability calculus—and every other constraint imposed by the correct epistemology—and X always maximizes expected utility. We will further assume that, when it comes to defending propositions, X's aim is to defend the truth. Thus, our task will be to say, for our idealized X, under what conditions X will defend a proposition.

Given the extent of this idealization, it should be clear that we will not end up with an accurate empirical theory of how actual persons to whom we are wont to apply the epithet 'rational' behave. But then this is not our aim. Ours is a normative enterprise whose aim is to construct part of a theory of valid criticism. It is not that, once we construct our theory, we will be in a position to indict any person who fails to satisfy the constraints the theory imposes upon X. After all, we can surely excuse those transgressions which are due to a person's failure to have X's logical and mathematical acumen (or, for that matter, a complete degree-of-confidence function). We can and should excuse an investigator for, say, accepting an inconsistent set of propositions whose inconsistency she, unlike X, does not see.

But, once we have our theory, we *will* be in a position to indict those persons whose violations of the theory's constraints are not excusable on grounds of failure of acumen (or absence of degree-of-confidence function). Thus we can and should indict the same investigator if, *after* having seen a reductio of the inconsistent set of propositions, she continues nonetheless to accept the set. In other words, having constructed a theory about our idealized X, we will be in a position to draw two morals concerning *actual* persons: (a) that, in order to construct a decisive criticism of an actual person's acceptance of a set of propositions, it suffices to show that she is thereby violating the constraints our theory places upon X; and (b) that, in order to undermine a criticism of an actual person's acceptance of a set of propositions, it suffices to show that she is thereby satisfying the constraints our theory places upon X. As with any normative theory, the plausibility of our theory of rational acceptance will turn, in the end, upon how it accords with our considered judgments.

III

We will approach our task in a slightly indirect way. Given that comprehensiveness is one of X's aims in deciding what to defend, it seems only natural that X's primary concern should focus upon determining what proposition constitutes the most comprehensive proposition she should defend—i.e., what her theory of the world should be or, as we will henceforth say, what she should *defend as strongest.* Accordingly, we will be seeking to determine first what proposition, S, X will defend as strongest. Having done that, we will then be able to say, for any proposition P, that X will defend P just in case S entails P.[9] Later we will see that we need not take quite so circuitous a route.

There is a rather natural way to begin. We know that X, being rational in the Bayesian sense, will choose to perform an act only if she has no option that bears greater expected utility. We know as well that, when her aim is to defend the truth, X has two desires that bear on her decision as to what to defend as strongest: the desire for truth and the desire for comprehensiveness. These two bits of information suggest the following strategy: represent the two desires in a utility function. If we do, we will be able to say that X's aim is to defend the truth just in case this *epistemic* utility function is her sole utility function. It will then follow from the fact that X is rational in the Bayesian sense and the fact that her aim is to defend the truth that she will defend P as strongest only if she has no other option that bears greater expected epistemic utility.

The strategy, as attractive as it is, will not work. It is not because there is any difficulty in representing the desire for truth in an epistemic utility function: we can let '1' represent the epistemic utility of defending P as strongest when P is true and let '-1' represent the epistemic utility of defending P as strongest when P is false. The trouble lies in trying to represent the desire for comprehensiveness in an epistemic utility function.

[9] Cf. Levi's use of 'accept as strongest' in *The Enterprise of Knowledge.* For the purpose of this section, we will view a theory of the world as excluding arithmetic and simply impose the requirement that X's theory of the world be consistent with the truths of arithmetic—a requirement we will assume X can meet. The fact that arithmetic is not finitely axiomatizable entails that there is no proposition that entails all the truths of arithmetic and, thus, that there is no proposition that could count as a comprehensive theory of arithmetic (as Burton Dreben reminded me). On the more direct route, alluded to in the next sentence and sketched in section IV, this special treatment of arithmetic becomes unnecessary.

In order to have a utility function that represents the desire for comprehensiveness—i.e., that represents the view that, all other things being equal, the more comprehensive a proposition is, the better off one is in defending it as strongest—we need a measure of comprehensiveness. That is, we need a function that, for each proposition, will spit out a number representing the degree of comprehensiveness that proposition bears. The rub is that, despite the efforts of a number of talented philosophers, no one has come anywhere close to describing a measure of comprehensiveness that will do the job. No measure described thus far has applied to more than a restricted domain of cases or a set of artificial languages so simple as to be of no general interest.[10] Our objective here is to give a global account of why a rational investigator would spurn skepticism and defend theories like those we commonly defend. The epistemic-utility approach will not, then, help us reach our objective.

That is not to say that the objective cannot be reached with a clear Bayesian conscience. It is no part of the Bayesian credo that we must be able to find utility functions by mere reflection upon our aims. If we can order propositions by the preferability for *X* of defending them as strongest when her aim is to defend the truth—with reflection upon our aims as a guide—we will have done all we need do.

But even this task requires that we be clear about what it is to harbor the aim to defend the truth. And, so far, one ingredient in that aim—the desire for comprehensiveness—remains obscure. Can we make sense of the desire for comprehensiveness once we admit that no one has yet found any measure of comprehensiveness? I suggest that we can. I suggest that we do not even need to assume the existence of any measure of comprehensiveness in order to say what the desire for comprehensiveness comes to. All we need be granted is the intelligibility of the stronger-than relation with which every philosopher is already familiar:

(D1) *P is stronger than Q* (*Q is weaker than P*) $=_{df}$ *P* entails *Q* and *Q* does not entail *P*.

The desire for comprehensiveness, I suggest, comes to this: a preference, all other things being equal, for defending *P* as strongest rather than *Q*, for any *P* and *Q* where *P* is stronger than *Q*.

[10] See Levi, *Gambling with Truth*, chaps. 4 and 5, for an example of the first sort of measure. Examples of the second kind abound, produced by such philosophers as Y. Bar-Hillel, R. Carnap, J. Hintikka, and K. Popper.

With this much done, we can take a first modest step toward solving the problem of rational acceptance. We can try to say how X, if rational, ought to adjudicate between the desire for truth and the desire for comprehensiveness when the choice is between defending P as strongest and defending Q as strongest, where P is stronger than Q. Saying this much will not, of course, suffice to give us a theory of rational acceptance. We will not yet have said how X ought to choose between rival theories neither of which entails the other. But it will do for a start.

As a first attempt, I suggest that X, if rational, will defend P as strongest rather than Q, when P is stronger than Q, just in case her confidence in the truth of what she defends as strongest will not thereby be reduced by an excessive degree. That is to say [where, for any P, 'prob(P)' represents X's degree of confidence that P],

(3) There is a number c greater than .5 such that, for any P and Q where P is stronger than Q, X, if rational, will prefer defending P as strongest to defending Q as strongest *if and only if* prob(P) is greater than $c \cdot \text{prob}(Q)$.

Some comments on (3).

As modest as (3) is, it may nonetheless seem on first blush to saddle us with an immodest task. (3) seems to entail that, once X is willing to defend the claim that the *Times* wrote that the president signed a bill, then, in order to determine whether X should also be willing to defend the claim that the president *did* sign the bill, we have to compute the degree of confidence X assigns to her entire theory (T) of the world, then compute the degree of confidence she assigns to that theory conjoined with the claim that the president did sign the bill (call this conjunct B) and *then* compute what fraction the second number is of the first. It would hardly be to the credit of (3) if it turned out that, in order to evaluate X's rather simple choice between defending T as strongest or (T & B) as strongest, we had to engage in so herculean a labor.

Fortunately, (3) requires no such thing. To tell whether defending (T & B) as strongest is preferable to defending T as strongest for X, we need know neither the value of prob(T) nor the value of prob(T & B). All we need know are the values of c and prob(B/T). According to (3), X is rational to prefer defending (T & B) as strongest to defending T as strongest if and only if

$$\text{prob}(T \,\&\, B) > (c \cdot \text{prob}(T))$$

Note that, since X is rational in the Bayesian sense and thus her degree-of-confidence function satisfies the axioms of the probability calculus,

$$(\text{prob}(B/T)\cdot\text{prob}(T)) > (c\cdot\text{prob}(T))$$

So X is rational to prefer defending (T & B) as strongest to defending T as strongest if and only if

$$\text{prob}(T \,\&\, B) = (\text{prob}(B/T)\cdot\text{prob}(T))$$

Dividing both sides by prob(T), that is to say if and only if

$$\text{prob}(B/T) > c$$

The value c takes will made a great difference to what X, if rational, will prefer to what. For any P and Q such that P is stronger than Q, the greater c's value for X, the greater prob(P) must be before she will prefer defending P as strongest—that is, the more skeptical X will be in defending propositions as strongest. Accordingly, the smaller the value of c for X, the less skeptical she will be.[11]

As one would expect, (3) prescribes that X's desire for comprehensiveness ought not to be untrammeled—c must be greater than .5. And, as one would expect, (3) prescribes that, at any time, the same value of c should determine all X's preferences. (X ought not selectively adopt special standards for certain propositions—say, for propositions she hopes are true.) (3) does not, however, prescribe any further constraints on what value c should take. This permissivism is deliberate. By allowing that c may take different values for different persons, we can explain how it is that some rational people *are* more skeptical than other rational people. In particular, relativizing c to persons (and times) allows us to explain why the skeptic is moved by an argument to the effect that it is logically possible that we are but brains in vats subject to constant illusion, yet we are not moved in the same way by that argument.

The effect the argument has on the skeptic's confidence that, say, she has limbs, is this: it makes her less than absolutely certain that she has any. But, even if the argument has the same effect on us, it does not move us to prefer to defend a tautology as strongest rather than the claim that we have limbs. For the value of c we nominate is such that defending the stronger proposition as strongest is still preferable for us to defending the weaker one. That is, we are will-

[11] Compare Levi's index of caution, q (*Gambling with Truth,* pp. 86-90).

ing to increase marginally our risk of defending a falsehood as strongest for the promise of greater comprehensiveness.

The skeptic is not.[12] The skeptic nominates a value of c equal to 1. So, no matter how little confidence the skeptic loses in the truth of the claim that she has limbs, c is high enough so that defending that claim as strongest is not preferable for her to defending a tautology. Nor is it any wonder that we get nowhere arguing with the skeptic and she nowhere with us. We have different epistemic tastes, and *de gustibus non disputandum est.*

But enough time has been spent on (3)'s virtues. It is time we noticed that (3) suffers a mortal vice. Although attractive as a sufficient condition, it is defective as a necessary condition. The reason is simple: we want our preference relation to be transitive. That is, we want our theory to say that, for any P, Q, and R, if X is rational and if X prefers defending P as strongest to defending Q as strongest and X prefers defending Q as strongest to defending R as strongest, then X should prefer defending P as strongest to defending R as strongest. (3) entails the contrary.

Consider an arbitrary P, Q, and R. Let P be stronger than Q and Q stronger than R. Let $\text{prob}(R) = 1$, $\text{prob}(Q) = .91$, and $\text{prob}(P) = .88$. And suppose that $c = .9$ for X. According to (3), X will, if rational, prefer defending P as strongest to defending Q as strongest and prefer defending Q as strongest to defending R as

[12] For example, Descartes writes, "Science is in its entirety true and evident cognition. He is no more learned who has doubts on many matters than the man who has never thought of them; nay he appears to be less learned if he has formed any wrong opinions on any particulars. Hence it were better not to study at all than to occupy one's self with objects of such difficulty, that, owing to our inability to distinguish true from false, we are forced to regard the doubtful as certain; for in these matters any hope of augmenting our knowledge is exceeded by the risk of diminishing it. Thus in accordance with the above maxim we reject all such merely probable knowledge and make it a rule to trust only what is completely known and incapable of being doubted." "Rules for the Direction of the Mind," in *Philosophical Works*, vol. 1, E. S. Haldane and G. R. T. Ross, trans. (Cambridge: University Press, 1911), p. 3.

Throughout this paper I am using the term 'skepticism' to refer to the doctrine that (a) no person is rational to defend P when her aim is to defend the truth unless she is certain P is true. (a) is to be distinguished from two other doctrines which can also be called skeptical: (b) no person is rational to be certain of anything; and (c) no person is rational, for any P, even to place more confidence in P's truth than she places in the truth of P's denial. I would claim that Descartes endorsed (a), assumed the denial of (c), and argued against (b). If this claim is correct, it is not hard to see why Descartes would (as he did) view his skeptical doubts as quite academic—devoid of immediate practical consequences. For these doubts, consisting as they did in the worry that, though (c) is false (b) might be true and, thus, no person might be rational to defend anything when her aim was to defend the truth, never questioned the propriety of the probabilistic reasoning that guides life.

strongest, but X will not prefer defending P as strongest to defending R as strongest.

Putting theory first and elaboration second, here is how I propose to remedy this defect. Consider the following definition:

(D2) For any P and Q and number c, a set of propositions A is a *good chain* from Q to P with respect to c for $X =_{df}$
(i) A is finite and linearly ordered;
(ii) A has Q for its first member and P for its last; and
(iii) for any two members of A, S and S', such that S' is the immediate successor of S, S' is stronger than S and $\text{prob}(S')$ is greater than $c \cdot \text{prob}(S)$.

The following replacement for (3) will do the trick:

(4) There is a number c greater than .5 such that, for any P and Q where P is stronger than Q, X, if rational, will prefer defending P as strongest to defending Q as strongest *if and only if* there is a good chain from Q to P with respect to c for X.

(D2) and (4) give us transitivity in a straightforward way: when it comes to defending propositions as strongest, if (and only if) (3) says X ought to prefer P to Q and says X ought to prefer Q to R, (4) says X ought to prefer P to R.

Note that it is not simply a general intuition concerning rational preference ordering which motivates the demand for transitivity.[13] A theory that says, with respect to some P and R where P is stronger than R, that X ought not to prefer to defend P as strongest rather than R given that her aim is to defend the truth, is saying that X, given R, ought not to infer P—that X does not have a good argument from R for P. And reflection on the nature of argument and rational inference will alone suffice to motivate the judgment that any theory that prohibits the inference from R to P, yet prescribes the inference from R to Q and the inference from Q to P—as (3) does—is inadequate.

Note, too, that we get more than just transitivity from (D2) and (4). What we get, as well, is an account of how, for some P and Q where P is stronger than Q, X may be rational to prefer defending P as strongest to defending Q as strongest when her aim is to defend the truth, even when X is absolutely certain that Q is true and extremely confident that P is false. By (D2) and (4), so long as there is a good chain (good argument) for X from Q to P with respect to the value of c she nominates, X will, if rational, prefer defending P

[13] Joan Weiner suggested the following motivation to me in conversation.

as strongest rather than *Q*—there is no need that prob(P) be greater than .5.

Unfortunately, we also get from (D2) and (4) something we do not want. Consider, once again, the case in which X is confronted by a fair one-thousand-ticket lottery. Assume, again, that X has the natural degree-of-confidence function with respect to claims concerning the outcome of this lottery. Let T be the proposition, 'There is a one-thousand-ticket lottery, and one and only one ticket will win'. For any number i, let 'Lt_i' represent the proposition, 'Ticket number i will lose'. Now consider the following ordered set of propositions: $\{T, (T \mathbin{\&} Lt_1), (T \mathbin{\&} Lt_1 \mathbin{\&} Lt_2), \ldots, (T \mathbin{\&} Lt_1 \mathbin{\&} Lt_2 \mathbin{\&} \ldots \mathbin{\&} Lt_{990})\}$. By (D2), this set constitutes a good chain from T to $(T \mathbin{\&} Lt_1 \mathbin{\&} Lt_2 \mathbin{\&} \ldots \mathbin{\&} Lt_{990})$ with respect to .9 for X. If X nominates the value .9 for c, then (4) states that X should prefer defending $(T \mathbin{\&} Lt_1 \mathbin{\&} Lt_2 \mathbin{\&} \ldots \mathbin{\&} Lt_{990})$ as strongest to defending T as strongest. But I doubt that anyone will want to say that, confronted with a fair one-thousand-ticket lottery, X has a good argument for the claim that one of the last ten tickets will win.

This unwanted consequence of (D2) and (4) leaves us with a two-part problem. One part of the problem is to say what special policy X ought to follow vis-à-vis predicting the outcome of the lottery. The other part of the problem is to say what features the case just outlined displays by virtue of which it deserves to be treated by a special policy. Notice that, unless we can solve this second part of our problem, any solution to the first part will be cold comfort. For cases like the lottery case are legion.

Consider, for example, the following case. X is a coroner. She is certain Harry died between midnight and one o'clock in the morning (call the claim 'H') but feels it equally likely that Harry died at one time in that interval as it is that he died in another. Divide the interval between midnight and 1 A.M. into 1000 subintervals of equal length. Number the intervals in order of their increasing proximity to 1 A.M. For any number i, let 'D_i' represent the proposition, 'Harry did not die in interval number i'. Consider now the following ordered set of propositions: $\{H, (H \mathbin{\&} D_1), (H \mathbin{\&} D_1 \mathbin{\&} D_2), \ldots, (H \mathbin{\&} D_1 \mathbin{\&} D_2 \mathbin{\&} \ldots \mathbin{\&} D_{990})\}$. By (D2), this is a good chain from H to $(H \mathbin{\&} D_1 \mathbin{\&} D_2 \mathbin{\&} \ldots \mathbin{\&} D_{990})$ for X with respect to .9. But no one will think that X, in this case, has a good argument from H to the claim that Harry died in the last thirty-six seconds before one o'clock.

The case of the coroner is every bit as disturbing as the case of the lottery—and, it would seem, for the same reason. But, since the cases are distinct (one about a time of death, one about the outcome of a lottery), a special policy for one case is not *ipso facto* a policy for the other. Therefore, constructing a policy for the lottery case alone will not leave us appreciably better off—we have the coroner case (and an infinite number of variations of it) in the wings. It will be worth while to construct a special policy for the lottery case only if we can provide a general account of when the policy is to be invoked (as in the lottery and coroner cases) and when it is not (as in the cases where there is a good chain from Q to P for X and prob(P/Q) is low, yet we think there is a good argument from Q to P).

What is it, then, about the lottery case which singles it out for special treatment? I suggest that it is an *arbitrariness* in the argument from T to $(T \,\&\, Lt_1 \,\&\, Lt_2 \,\&\, \ldots \,\&\, Lt_{990})$—in fact, in each step of the argument: each member of the chain from T to $(T \,\&\, Lt_1 \,\&\, Lt_2 \,\&\, \ldots \,\&\, Lt_{990})$ is arbitrary with respect to its predecessor, given the value of c for X. And in what does this arbitrariness consist? I suggest, where $\approx$ means "is roughly equal to":

(D3) For any P, Q, and number c, P is *arbitrary* with respect to Q and c for $X =_{df}$ P is a member of a set of propositions $\{R_1, R_2, \ldots, R_n\}$ such that
- (i) each R_i is stronger than Q;
- (ii) for each R_i, prob(R_i) is greater than $c \cdot$prob(Q);
- (iii) for each i and j, $i \neq j$, prob(R_i/R_j) $\approx$ prob(R_j/R_i) $\not\approx$ prob(R_i);
- (iv) prob($R_1 \,\&\, R_2 \,\&\, \ldots \,\&\, R_n$) is not greater than $c \cdot$prob(Q); and
- (v) there is no set A satisfying (i)-(iv) such that each member of A is both stronger than at least one R_i and stochastically independent of P.

Thus $(T \,\&\, Lt_1)$, by virtue of its membership in $\{(T \,\&\, Lt_1), (T \,\&\, Lt_2), \ldots, (T \,\&\, Lt_{1000})\}$, is arbitrary with respect to T and .9 for X. Similarly, $(T \,\&\, Lt_1 \,\&\, Lt_2)$, by virtue of its membership in $\{(T \,\&\, Lt_1 \,\&\, Lt_2), (T \,\&\, Lt_1 \,\&\, Lt_3), \ldots, (T \,\&\, Lt_1 \,\&\, Lt_{1000})\}$, is arbitrary with respect to $(T \,\&\, Lt_1)$ and .9 for X—and so on. The motivation for clauses (i) and (ii) and most of clause (iii) should be clear. The rest, however, deserves some comment.

First, one might ask why (v) is included. The reason is that, without (v), (D3) would be vulnerable to the following problem.[14] Let P be any proposition stronger than T but otherwise unrelated

[14] Pointed out to me by Allan Gibbard. I am indebted to Gibbard for criticism of earlier versions of (D3).

to the lottery such that prob(P) happens to be equal to 1000/1001. Now consider the following set of propositions: $\{P, (T \,\&\, (\sim P \vee Lt_1)), (T \,\&\, (\sim P \vee Lt_2)), \ldots, (T \,\&\, (\sim P \vee Lt_{1000}))\}$. This set satisfies (i)-(iv) in (D3),[15] but surely we do not want to allow this fact to count against P. (v) saves P from this undeserved epistemic taint. The set just described has been cooked up with a more familiar set in mind: $\{(T \,\&\, Lt_1), (T \,\&\, Lt_2), \ldots, (T \,\&\, Lt_{1000})\}$. And this more familiar set is such that it satisfies (i)-(iv), each of its members is stronger than at least one member of the cooked-up set, and each of its members is stochastically independent of P; i.e., for every i, prob($P/(T \,\&\, Lt_i)$) = prob(P) and prob($(T \,\&\, Lt_i)/P$) = prob($T \,\&\, Lt_i$).

Second, one might wonder why (iv) does not impose the stronger condition that the set of R_is be inconsistent—or the even stronger condition that prob($R_1 \,\&\, R_2 \,\&\, \ldots \,\&\, R_n$) = 0. To see why these conditions are not imposed, let us return to the lottery case.[16] Suppose that X is not certain of what day the lottery will be held, but has a degree of confidence equal to .5 that it will be held on May Day. For any i, let $\sim Wt_i$ be 'ticket number i will not win on May Day'. Let Q be any proposition unrelated to the lottery such that prob(Q) $\neq 0$. Consider now the following ordered set of propositions: $\{Q, (Q \,\&\, \sim Wt_1), (Q \,\&\, \sim Wt_1 \,\&\, \sim Wt_2), \ldots, (Q \,\&\, \sim Wt_1 \,\&\, \sim Wt_2 \,\&\, \ldots \,\&\, \sim Wt_{990})\}$. By (D2), this set constitutes a good chain from Q to $(Q \,\&\, \sim Wt_1 \,\&\, \sim Wt_2 \,\&\, \ldots \,\&\, \sim Wt_{990})$. And that chain no more represents a good argument than does the chain from T to 'one of the last ten tickets will win'. Thus we expect of (D3) that it will treat these chains alike—that it will say, for example, that $(Q \,\&\, \sim Wt_1)$, by virtue of its membership in $\{(Q \,\&\, \sim Wt_1), (Q \,\&\, \sim Wt_2), \ldots, (Q \,\&\, \sim Wt_{1000})\}$ (call this set F), is arbitrary with respect to Q and .9 for X. And (D3) does indeed do the job. But notice that (D3) would not be doing our bidding were (iv) strengthened in either of the ways con-

[15] That the set satisfies (iii) is perhaps hardest to see. So let us show why, for any i, prob($P/(T \,\&\, (\sim P \vee Lt_i))$) = prob($(T \,\&\, (\sim P \vee Lt_i))/P$). Since, by the multiplication axiom of the probability calculus, the left side is equal to prob($P \,\&\, T \,\&\, (\sim P \vee Lt_i)$)/prob($T \,\&\, (\sim P \vee Lt_i)$) and the right side is equal to prob($P \,\&\, T \,\&\, (\sim P \vee Lt_i)$)/prob($P$) and since these fractions have a common numerator, it will suffice to show that their denominators are equal. Prob(P) = prob($T \,\&\, P$), since P entails T. So it will do to show that prob($T \,\&\, (\sim P \vee Lt_i)$ = prob($T \,\&\, P$); i.e., by the multiplication axiom, that [prob(T)·prob($(\sim P \vee Lt_i)/T$)] = [prob(T)·prob(P/T)]; i.e., dividing both sides by prob(T), that prob($(\sim P \vee Lt_i)/T$) = prob(P/T); i.e., that prob($(\sim P \vee Lt_i)/T$) = 1000/1001. And to show this we need notice only that (by the probability calculus) prob($(\sim P \vee Lt_i)/T$) = [1 − prob($\sim(\sim P \vee Lt_i)/T$)] = (by the propositional calculus) [1 − prob($(P \,\&\, \sim Lt_i)/T$)] = (by the multiplication axiom) [1 − (prob(P/T)·prob($\sim Lt_i/(P \,\&\, T)$))] = [1 − (1000/1001 · 1/1000)] = 1000/1001.

[16] John Carriero brought the following kind of case to my attention.

sidered. For F is not inconsistent—after all it is not inconsistent that the lottery will not be held on May Day and thus, not inconsistent that no ticket will win on May Day. Moreover, X is not certain that the conjunction of F's members is false, insofar as X is not certain that the lottery will be held on May Day.[17]

Third, one might question the use of '$\approx$' in (iii) rather than '$=$'. To see the rationale, suppose X is confronted with a lottery exactly like the one described earlier in all but one respect: as i approaches 1000, X is ever so slightly more confident that t_i will lose, so that, for every i greater than 1, prob(T & Lt_i) is greater than prob(T & Lt_{i-1}) by an arbitrarily small amount. For example, X may be certain that the winning ticket will be drawn from a drum and that, for any i greater than 1, t_i is just slightly lighter than t_{i-1} and, hence, that much more likely to land on top. Were '$=$' substituted for '$\approx$' in (iii), then members of the set of propositions, $\{(T \& Lt_1), (T \& Lt_2), \ldots, (T \& Lt_{1000})\}$, would *ipso facto* turn out not to be arbitrary with respect to T and .9 for X in this case. But I do not think that so slight a change in X's degree-of-confidence function ought to change our attitude toward the propriety of X's argument from T to 'T and one of the last ten tickets will win'.[18]

What singles out the good chain from T to $(T \& Lt_1 \& Lt_2, \& \ldots \& Lt_{990})$ in the lottery case and the good chain from H to $(H \& D_1 \& D_2 \& \ldots \& D_{990})$ in the coroner case is, then, the fact that each of these chains is *tarnished*, where

(D4) For any P and Q, a good chain for X from Q to P with respect to c is *tarnished* $=_{df}$ there are at least two propositions, S and S', in the chain such that S' is a successor of S and S' is arbitrary with respect to S and c for X.

This brings us to the other part of our problem. What special treatment ought these tarnished good chains receive? I suggest that no investigator ought, when her aim is to defend the truth, defend any prediction stronger than T concerning the outcome of the lottery or any claim stronger than H concerning the time of Harry's death. That is, I suggest that no tarnished good chain represents a

[17] Notice that, once X's degree of confidence that the lottery will not be held on May Day exceeds c, the members of F cease to be arbitrary with respect to Q and c. But of course in the event, $\{Q, (Q \&$ no ticket will win on May Day$)\}$ is a good chain for X—i.e., X can, given her degree of confidence that the lottery will not be held on May Day, move directly without intermediate steps from Q to the prediction that Q and no ticket will win on May Day.

[18] It would be nice to do better than merely substitute '$\approx$' for '$=$', but I am afraid I do not know how to effect more precision. I am, however, somewhat consoled by my conviction that this imprecision is mirrored quite faithfully by the actual state of our art of rational criticism.

good argument. Hence, the following successor to (4):

(5) There is a number c greater than .5 such that, for any P and Q where P is stronger than Q, X, if rational, will prefer defending P as strongest to defending Q as strongest *if and only if* there is an untarnished good chain from Q to P with respect to c for X.

But, although (5) does take care of the lottery case and its ilk, it will not quite do.

In disposing of the epistemic-utility strategy earlier in this section, I said that our objective here was to say how X could be rational to defend the sort of theories we commonly defend and *not* be a skeptic—where, by 'a skeptic', I meant a person who will not defend any proposition of whose truth she is not certain. (5) has a role to play in reaching that objective. Its role is to say how X could be rational to prefer to defend, for example, the historian's theory T as strongest rather than a proposition of which she is certain—say, a tautology.

(5) fills that role. (5) says that, as long (and just as long) as there is an untarnished good chain from $(P \vee {\sim}P)$ to T for X with respect to the value of c that X nominates, X should prefer defending T as strongest to defending $(P \vee {\sim}P)$. But there is something wrong with the way (5) fills its role: according to (5), even if there is an untarnished good chain from $(P \vee {\sim}P)$ to T with respect to c for X—even if, given only a tautology, X has a good argument for T—defending T as strongest may still fail to be preferable for X to defending a proposition that is weaker than T.

Suppose there is an untarnished good chain from $(P \vee {\sim}P)$ to T with respect to c for X. Suppose Q is the penultimate member of the chain and that T, the ultimate member, is logically equivalent to $(Q \,\&\, R)$. Prob$((Q \,\&\, R)/Q)$ is then greater than c. But notice that it is compatible with the fact that prob$((Q \,\&\, R)/Q)$ is greater than c that prob$((Q \,\&\, R)/R)$ is *not* greater than c. In fact, it is compatible with the fact that prob$((Q \,\&\, R)/Q)$ is greater than c that there is no untarnished good chain from R to $(Q \,\&\, R)$ with respect to c for X. And if there is no such chain, (5) dictates that, although defending T as strongest is preferable for X to defending $(P \vee {\sim}P)$ as strongest, it is not preferable to defending R as strongest.

In other words, in order to show that X should prefer defending T as strongest to defending R as strongest, it is not enough to show that T is logically equivalent to $(Q \,\&\, R)$ and that there is a good argument from $(P \vee {\sim}P)$ to Q and from Q to $(Q \,\&\, R)$ with respect to c for X. According to (5), one needs to show that from R alone one can find yet another good argument for T. In fact, (5) dictates that, for each proposition that *T* entails, one needs to find an argument from that proposition alone to *T* in order to show that defending *T* as strongest is preferable to defending anything weaker. Given the existence of a good argument for *T* from $(P \vee \sim P)$—the weakest proposition that *T* entails and a proposition of whose truth *X* is certain—this is a draconian requirement. Thus the following revision of (5):

(P1) There is a number *c* greater than .5 such that, for any *P* and *Q* where *P* is stronger than *Q*, *X*, if rational, will prefer defending *P* as strongest to defending *Q* as strongest *if and only if* there is for *X* an untarnished good chain to *P* with respect to *c* either from *Q* or from $(P \vee \sim P)$.[19]

We can now finish with the preferability relation as defined for propositions related by the stronger-than relation. All we need do is add to (P1) the following claim, which requires no comment:

(P2) For any *P* and *Q* where *P* is stronger than *Q*, *if X*, if rational, will not prefer defending *P* as strongest to defending *Q* as strongest, *then X*, if rational, will prefer defending *Q* as strongest to defending *P* as strongest.

(P1) and (P2) do not, of course, tell us what *X* should defend as strongest. They do not say how *X* ought to adjudicate between two propositions neither of which entails the other. But, if (P1) and (P2) do not tell us what ought to be *X*'s theory of the world, they do give us a clue as to where we should look for it.

We know that, in order for a proposition *T* to be the theory of the world *X* ought to defend, defending *T* as strongest will have to be preferable for *X* to defending anything weaker. Otherwise *X* will, by (P2), prefer defending some weaker proposition as strongest and will thus be required, on pain of irrationality, not to defend *T*. We also know that defending *T* as strongest will have to be preferable for *X* to defending anything stronger. Otherwise *X* will prefer to defend the stronger proposition as strongest and will thus be required, on pain of irrationality, not to defend *T* as strongest. This suggests that, to find the theory of the world *X* ought to defend, we should look for a proposition that is *nice* for *X*, where

(D5) *P* is *nice* for $X =_{df}$ with respect to the value of *c* for *X*, there is for *X* an untarnished good chain from $(P \vee \sim P)$ to *P* and no untarnished good chain from *P* to any proposition stronger than *P*.

[19] And *Q* is not super. See (D6) below.

This may not, however, suffice to end our quest for the theory of the world X ought to defend. We may find that there is more than one proposition which is nice for X. For example, $(T \& P)$, $(T \& Q)$, and $(T \& R)$ may all be nice for X, because of the absence of a good chain from any of them to their conjunction—or even to the conjunction of two of them. What proposition ought then to be X's theory of the world?

The criterion of credibility will serve to eliminate some pretenders. Having given the desire for comprehensiveness its way in constructing nice propositions, we can submit to the demand for truth: P, even though nice for X, ought *not* to be in contention for X's theory of the world if there is another nice proposition Q such that prob(Q) is greater than prob(P). Of course, even after we eliminate all but the most credible of the nice propositions, we may be left with more than one: there may be a tie for most probable nice proposition. Thus, to return to our example, if $\text{prob}(T \& Q) > \text{prob}(T \& R)$ and $\text{prob}(T \& P) = \text{prob}(T \& Q)$, the criterion of credibility will eliminate only $(T \& R)$. If so, the desire for truth will once again provide good guidance: most would agree that X ought, in the face of two equally good theories of the world, hedge her bets. She should conclude that at least one of them, she knows not which, is correct. That is, X ought to defend their disjunction as strongest—in our example, $(T \& (P \vee Q))$.

Let us call the theory of the world that X ought to defend *super* for X. Then we can say that

(D6) P is *super* for $X =_{\text{df}} P$ is the disjunction of each proposition Q such that, for X, Q is nice and there is no nice proposition R such that $\text{prob}(R) > \text{prob}(Q)$.

Having thus found what proposition X will defend as strongest, we can find out, for *any* proposition P, whether X will defend P:

(P3) For any P, X, if rational, will defend P *if and only if* P is entailed by that proposition which is super for X.

(P3) completes our constructive task.[20] In telling us what X, if rational, will defend on the assumption that X's aim is to defend the truth, (P3) tells us what X would, if rational, defend were her aim to defend the truth—i.e., what X, if rational, accepts. It remains only to reflect briefly on some of the theory's features.

[20] I have not, of course, given a complete preference ordering for X over acts of defending propositions as strongest. I have only displayed enough of such an ordering to accomplish my purpose: to construct a theory that will say, as (P3) does, when X is rational to defend a proposition given that her aim is to defend the truth.

IV

Before I began constructing the theory in the last section, I noted that there are a number of prescriptions that, given our endorsement of (2), we should expect a theory of rational acceptance to make. This, in effect, imposed a number of conditions of adequacy on the theory just constructed. Accordingly, the first item on the agenda is to show that this theory satisfies those conditions.

The first condition is that the theory satisfy (2), the canon of rationality that forbids rational persons to ignore reductio arguments. It will be obvious from (P3) that the theory satisfies the first two clauses in (2): if X is rational X will defend the conjunction and consequences of any proposition she will defend. Though this may be a little less obvious, the theory also satisfies the third clause in (2): if rational, X will not defend a contradiction. For suppose otherwise. Then it is possible for there to be a proposition P such that P is a contradiction and P is entailed by V, the proposition that is super for X. Thus, by the theory, there is, with respect to some c greater than .5, an untarnished good chain from $(P \vee \sim P)$ to V for X. Since X is rational in the Bayesian sense, and thus her degree-of-confidence function satisfies the axioms of the probability calculus, $\text{prob}(P \vee \sim P) = 1$, $\text{prob}(P) = 0$, and, since V entails P, $\text{prob}(V) = 0$. But note that, for there to be an untarnished good chain from $(P \vee \sim P)$ to V for X with respect to some c greater than .5, where $\text{prob}(P \vee \sim P) = 1$ and $\text{prob}(V) = 0$, there have to be two propositions in the chain S and S' such that S' is the immediate successor of S, $\text{prob}(S) > 0$, $\text{prob}(S') = 0$, and $\text{prob}(S') > (c \cdot \text{prob}(S))$, where c is positive. And, by arithmetic, that is impossible.

The second condition our theory must meet is that it tell us that not all the predictions are to be defended in the lottery case. The theory does—dictating that there will be no ticket of which X is rational to predict that it will lose. Even though X is willing to defend T ('There is a one-thousand-ticket lottery in which one and only one ticket will win'), there is no i for which there is an untarnished good chain from T or anything else X is willing to defend to $(T \,\&\, Lt_i)$. Thus for no i is there a nice proposition for X which entails $(T \,\&\, Lt_i)$. And, since no proposition can be super for X unless it is entailed by a proposition that is nice for X, there is no i such that $(T \,\&\, Lt_i)$ is entailed by the proposition that is super for X.

The third condition on the theory of rational acceptance is that it account for the willingness of a rational person (such as our historian) to defend certain—yet not all—propositions of whose truth she is not confident. The theory meets this condition as well. So long (and, of course, only so long) as the conjunction T of all the propositions expressed in her book is entailed by the proposition S which is super for her, the historian (now imagined as a rational person in the Bayesian sense) will, if rational, defend that conjunction. The historian need not invest much confidence in T for this condition to be met—any more than she need invest much confidence in the proposition that is super for her. The argument from tautology to S (and thence to T which S entails) involves a measured sacrifice of the desire for truth in exchange for satisfying the desire for comprehensiveness.

But note that, as freely as the theory licenses the sacrifice of truth for comprehensiveness, allowing the historian to defend a theory she is confident is false, it does not license wanton sacrifice. To be worthy of defense, a proposition must be entailed by the most credible of the nice propositions—propositions to which there is an untarnished good chain from $(P \vee {\sim}P)$ with respect to c for X. And this is a condition which no garden-variety implausibility will meet.

Now to the second matter on the agenda. At the beginning of the last section I said that we would take an indirect route in determining what propositions X ought to defend: we would determine first what theory of the world X ought to defend and then conclude that, for any P, X ought to defend P just in case P is entailed by that theory. But, as I suggested then, we need not take so circuitous a route. It is time to expand a bit upon that suggestion.

In virtually every case in which, for some P, a question concerning the acceptability of P arises, context limits the ramifications of the question. Rather than raise the question, "What theory of the world should X accept?" our concern about the acceptability of P typically raises, at most, the question "What theory of the subject matter relevant to P should X accept?" And, even then, there will generally be a considerable body of theory, M, such that all investigators take for granted that, whatever theory of the subject matter X ought to accept (and whether or not this theory entails P), it ought to entail M. Thus our asking whether X ought, for example, to defend the claim that Bacon wrote *Hamlet* need not be an occasion to canvass candidates for X's theory of the world. We may indeed consider the acceptability of broader theories concerning Bacon's life and character, but we can grant that Bacon could write English, and we can leave quantum theory out of it entirely.

Although our theory was constructed without explicitly considering the role of context in localizing criticism, the theory is easily adapted to the local context. Instead of determining whether X should accept P by determining what theory of the world X ought to accept, we can determine whether X should accept P by determining what theory of the subject matter relevant to P X should accept. And we can say that X ought to accept P just in case P is entailed by the proposition about this "small world" which X ought to defend as strongest.[21]

Finally, some closing remarks.

I suggested earlier that our preoccupation with the propriety of what people defend as if their aim were to defend the truth indicates that a theory of rational acceptance, in the sense of 'acceptance' being employed here, deserves a prominent place in a theory of rational inquiry. Nonetheless, it should be noted that, in our sense of 'acceptance', a theory of rational acceptance occupies a less prominent place in epistemology than it is typically accorded. For, although X's degree-of-confidence function (together with her choice of c) determines what she will accept if she is rational, the contour of that function is presumably in no way determined by what she accepts—after all, acceptance is not a state of confidence. Indeed (as Bayesians have been saying for a long time) the traditional problems of epistemology—e.g., the problem of induction, the controversy over foundationalism—arise, not for the theory of rational acceptance proper, but rather for the theory of rational degree-of-confidence functions.

Some may feel that the theory of rational acceptance should turn out to be more important than the view in this paper makes it out to be. It is surely an understandable sentiment. But those who harbor it should remember that it is a sentiment which cannot be cheaply indulged. For anyone who would accuse the present view of unjustly belittling acceptance must tell us how she proposes to meet the Bayesian challenge: she must tell us how the acceptance of propositions could possibly play a more central role in rational human conduct.

MARK KAPLAN

University of Wisconsin, Milwaukee
Southern Methodist University

[21] I recognize that one wants to hear more about local contexts than I have said here. I am not sure how much more there is to be said. But I am sure that, whatever there is to be said, I haven't the space to say it here. Parenthetically, the term 'local context', as well as my concern with the phenomenon it denotes, is due to Isaac Levi. The expression 'small world' is lifted from Savage, *op. cit.*, p. 82.