On the Relation Between Categorical and Probabilistic Belief
Author(s): Daniel Hunter
Source: *Noûs*, Vol. 30, No. 1 (Mar., 1996), pp. 75-98
Published by: Wiley
Stable URL: http://www.jstor.org/stable/2216304
Accessed: 05-10-2017 02:11 UTC

NOÛS 30:1 (1996) 75–98

*On the Relation between Categorical and Probabilistic Belief**

DANIEL HUNTER

Introduction

Belief is sometimes taken to be an all-or-nothing affair. One either believes something or one doesn't, on this view of belief. Call belief so understood "categorical belief." Having a categorical belief is very much like having a penny in one's pocket: the penny is either in your pocket or it isn't—there isn't any in-between. Similarly, when belief is understood categorically, one either has a particular belief or one does not have it—there is no in-between. In the categorical sense of belief, we may represent the belief state of an individual simply by a set of propositions, namely those propositions believed by that individual. Representing the belief state of an individual by a set of propositions makes clear why belief so construed is categorical, for a given proposition either belongs to the set or it doesn't. Set membership is not a matter of degree, at least within traditional set theory.

At other times, however, we wish to think of belief as a matter of degree, not as an all-or-nothing matter. We recognize that there are some things we believe more strongly than others. For example, I believe that $2 + 2 = 4$ and I also believe that Neanderthals buried their dead, but I believe the former proposition much more firmly than the latter. If beliefs can differ in strength, then we might attempt to quantify just how much stronger one belief is than another. This might lead to assigning numerical values to beliefs, representing just how strongly each is believed. Indeed, the Bayesian theory of subjective probability is an attempt to do just that. Bayesian probability theory has much to be said in its favor, yet it still remains controversial as a normative account of belief and especially as a descriptive account of human belief. I will not enter into this controversy here, but will instead take for granted that the Bayesian theory characterizes a kind of belief which, to distinguish it from all-or-nothing categorical belief, will be called "probabilistic belief."

An interesting question arises: What exactly is the relation between categorical and probabilistic belief? Are they entirely different notions or is one definable in terms of the other? Reducing probabilistic belief to categorical belief seems impossible on the face of it, since the notion of categorical belief contains much less structure than that of probabilistic belief. It is entirely unclear how, simply given a set of beliefs, we could derive the precise numerical values required for probabilistic belief. A more promising route would be to define categorical belief in terms of probabilistic belief—perhaps by saying that a categorical belief is one that is probable to degree θ or greater, where θ is some number sufficiently close to one. Some have argued, though, that because of the Lottery Paradox (to be discussed later) and related examples, no such reduction is possible. If they are right, then there is no clear way of reducing one type of belief to the other.

This situation is unsatisfactory in a number of respects. For one thing, it requires that we develop separate logics or formal theories to account for reasoning with beliefs of each kind. A unified account would be preferable. For another, it leaves mysterious what these two kinds of belief might have in common that would warrant calling each of them a kind of "belief."

This paper tries to show that there is a closer connection between categorical and probabilistic belief than some have thought. A wide variety of reasons have been offered for thinking these two notions distinct, but I want to focus on the attempt to pinpoint principles of reasoning that apply to one type of belief but not the other. Harman (1986) and Stalnaker (1987) have developed detailed accounts of reasoning and belief revision based on principles violated by probabilistic belief. They therefore argue against reducing categorical belief to probabilistic belief. I show that their arguments fail when we analyze categorical belief as a vague, or fuzzy, notion. But I will also argue that some intuitively plausible principles of reasoning, though strictly speaking incompatible with a probabilistic analysis of belief, nonetheless turn out to be in a certain sense "close" to the truth under such an analysis.

Problems for the Reduction of Categorical to Probabilistic Belief

One might think that there is an easy way to define categorical belief in terms of probabilistic belief. Specifically, we might be tempted to think that categorical belief is just probabilistic belief to a sufficiently high degree—that x categorically believes a proposition P iff x's subjective probability for P is greater than or equal to θ, where θ is some number less than, but sufficiently close to, one.

A number of objections to this simple proposal have been given. It is claimed, for example, that categorical belief is closed under deduction—if something follows from one's beliefs then that thing is or should be believed (or should be believed if it is recognized that it follows from one's beliefs—the niceties of the formulation are not too important now). However, high probability is not closed under deduction. One can have a number of statements each of which has a very high probability while some consequence of those statements does not have a high probability (and may even have zero probability). The classic example of this situation is Kyburg's Lottery Paradox (Kyburg 1961, 1970): there is a lottery with a large number of tickets, say, greater than a million. The lottery is fair, so for each ticket the probability that it is *not* the winning ticket is very high. Let L be the set of all statements of the form "Ticket i will not win" for i from 1 to the number of tickets. Then it is a consequence of L that no ticket is the winning ticket. But, we may suppose, it is highly probable, perhaps even certain, that some ticket is the winning ticket. Hence we have a set of statements, each of which is individually highly probable, that entails an improbable statement.

Harman (1986) concludes from the lottery example that if someone believes that there is a winning ticket, then she does not actually believe of each ticket that it is not the winning ticket but only believes that that is highly likely. Harman's view seems to be that one can take something to be a probable as one likes without actually believing it (as long as the probability is less than unity).

Stalnaker (1987) gives essentially the same response to the lottery example as does Harman. He does so in the course of defending the *Conjunction Principle* for beliefs, which says that if you believe P and you believe Q, then you ought to believe P&Q. The Conjunction Principle may appear much weaker than the principle that one's beliefs should be closed under entailment, but the combination of the Conjunction Principle with the principle that whatever is entailed by a single belief should also be believed, is equivalent to Deductive Closure for beliefs (assuming compactness: if a set propositions entails P, then some finite subset entails P). In the case of the lottery example, the assumption that each statement of the form "Ticket i will not win" is believed leads, by repeated application of the Conjunction Principle, to the false conclusion that it is (or should be) believed that no ticket will win. Again, we seem to face the choice of either rejecting a plausible principle of reasoning for categorical beliefs or else giving up the idea that categorical belief can be defined in terms of probabilistic belief. Harman and Stalnaker opt for the latter.

Spohn (1988), by way of motivating his development of a non-probabilistic system of belief revision, also cites the lottery example as a reason why categorical belief cannot be reduced to probabilistic belief.

Another principle that Harman thinks applies to categorical belief but not to probabilistic belief is the principle that a recognized inconsistency in one's beliefs is a reason to modify those beliefs. Normally, if you make a series of assertions and someone shows you that they are jointly inconsistent, then you have a problem. If you want to be rational, then you need to figure out some way of modifying your beliefs to restore consistency. However, there seems nothing wrong with assigning a high probability to propositions that are jointly inconsistent. Again, in the lottery example, each proposition of the form "Ticket i will not win" is highly probable as is the proposition "Some ticket will win." Taken together these propositions are inconsistent, but that fact is surely not a reason to modify the probabilities.

Many other authors (e.g. Levi (1967) and Lehrer (1980)) have advocated the view that it is a condition on rationality that one's beliefs be consistent. Some agree with Harman in denying that inconsistency *per se* is a mark of irrationality; they would deem an agent with inconsistent beliefs irrational only if the agent *recognizes* that the beliefs are inconsistent and does nothing to restore consistency. In this paper I will not worry about whether the agent recognizes his or her beliefs to be inconsistent. I will formulate the consistency condition simply as the claim that one's total set of beliefs should be consistent. Nothing in this paper, however, turns on whether we use this formulation or Harman's more refined formulation, since in the examples of inconsistent beliefs to be considered, we may, if we wish, assume that the agent is aware of the inconsistency.

The principle that one's beliefs should be jointly consistent should be distinguished from the principle that each individual belief should be consistent. Call the former principle "Global Consistency" and the latter "Local Consistency." [1] Global Consistency is a stronger condition than Local Consistency, yet if the Conjunction Principle is true and entailment is compact, then these two notions of consistency are equivalent. For if the set of all beliefs is inconsistent, then by compactness some finite set of beliefs entails a contradiction, and by the Conjunction Principle the conjunction of the members of that finite set is an inconsistent belief.

The three principles considered so far—Deductive Closure, Global Consistency, and the Conjunction Principle—are *static* principles of belief, principles that say how one's beliefs should be at a fixed moment in time. In contrast to static principles, there are *dynamic* principles for how beliefs should change over time, as new information comes in that requires modification of existing beliefs. Principles of this sort will be called principles of *belief revision*. Recent work on the revision of categorical beliefs (Spohn (1988), Gärdenfors (1988)) has yielded principles that can also be used to show a difference between categorical and probabilistic belief. An important principle in this category is that of Rational Monotony:

(RM) If P is believed and $\neg$Q is not believed, then P will still be believed if Q comes to be believed.

Here the condition "Q comes to be believed" is to be understood as "the only thing that is learned is that Q is true." (RM) is an axiom in Gärdenfors' (1988) theory of belief revision and is a consequence of Spohn's system for belief revision. (RM) is not correct if belief is equated with high probability and if the epistemic effect of learning Q can at least sometimes be represented by Bayesian conditionalization on Q. For example, if $\Pr(Q) = 0.5$, $\Pr(P \mid Q) = 1 - 2\Delta$, and $\Pr(P \mid \neg Q) = 1$, then $\Pr(P) = 1 - \Delta$. If $1 - \Delta$ is the threshold for belief (where Δ is some number close to zero), then we will have P believed and $\neg$Q not believed, but P will not be believed if Q is learned with certainty for its probability conditional on Q is below the threshold for belief. Hence if (RM) is correct when applied to categorical belief, then defining categorical belief in terms of a probabilistic threshold won't work.

We have, then, a wide array of plausible principles for categorical belief, none of which is satisfied if we try to define belief in P as P's subjective probability being at least $1 - \Delta$, for some small Δ. What are we to make of this? Should we reconcile ourselves to there being these two quite distinct forms of belief, neither or which is reducible to the other? I think that would be premature.

Deductive Closure & Co. Scrutinized

There is something extremely troubling about the use of the lottery example to drive a wedge between categorical and probabilistic belief. To begin with, it is useful to keep in mind that the lottery example was introduced by Kyburg as a counterexample to the Conjunction Principle and as a means to block the derivation of Deductive Closure and the Global Consistency Principle from weaker principles via the Conjunction Principle. Hence it is a bit cavalier to take the lottery example as a direct counterexample to a probabilistic rule for categorical belief. It is really only a counterexample to the combination of such a probabilistic rule with the Conjunction Principle, and we must decide which of these to give up on independent grounds.

If we do accept the principles advocated by Harman and Stalnaker, however, then the only reasonable conclusion about the lottery example is that we do not, if we are rational, believe of each ticket that it will lose. They grant that it is perfectly rational to believe of an individual ticket that it is highly probable that it will lose, but this does not translate, on their view, into the belief that it will lose. Thus the relevant belief is a belief about something's having a high probability, not a belief that itself is characterized probabilistically. It seems to me that this line just won't wash.

Harman and Stalnaker's approach threatens to destroy even the simplest of beliefs. Go back to November 1992, the morning of the presidential election. Very recent polls had shown Clinton with a sizable lead over Bush, leading pundits had predicted that Clinton would win, the events of the preceding weekend had reinforced the sense of a faltering, out-of-focus Bush campaign. In short, all the signs pointed toward a Clinton victory. It is now the morning of the election. Being a faithful listener to NPR and the MacNeil-Lehrer hour, I am well aware of all these signs the morning of the election. I believe, on that crisp November morning, that Clinton will be elected. I also believe the proposition that it is highly probable that Clinton will be elected but I don't just believe this proposition, I also believe that Clinton will be elected.

But now consider all the specific ways in which Clinton might win. He might win by exactly a million votes or he might win by exactly one million minus one votes or . . . or he might win by exactly one vote (here I am simplifying things by ignoring the electoral college and supposing that winning means winning the popular vote.) Suppose I believe he will not win by more than a million votes. There are therefore exactly a million possibilities compatible with what I believe for the number of votes by which Clinton will beat Bush. By Deductive Closure or by Stalnaker's Conjunction Principle, I cannot disbelieve (believe the negation of) each of these possibilities, for otherwise I would believe that Clinton will not win after all. Hence if my beliefs satisfy Deductive Closure or the Conjunction Principle, then there must be some proposition of the form "Clinton will not win by exactly i votes" which I fail to believe.

There's something funny about this, however. Suppose, for example, that I fail to believe the proposition that Clinton will not win by exactly twenty votes. (The choice of a particular number does not seem crucial.) This proposition seems a lot more certain than the proposition that Clinton will win. After all, polls have been misleading (remember Truman's victory over Dewey?), pundits are often wrong, and something could happen on election day itself that would drastically alter the outcome. On the morning of the election, a bet that Clinton will not win by exactly twenty votes would be a much safer bet than the bet that he will win. But how could the proposition that Clinton will not win by exactly twenty votes be more certain than the proposition that Clinton will win if I believe the latter but not the former?[2]

The general problem is that for almost any contingent belief you pick, you can partition the belief into numerous possibilities each of which is so unlikely that its negation is more believable than the original belief. By the Conjunction Principle or by Deductive Closure, your beliefs will be consistent only if you suspend judgment regarding at least some of these possibilities. But then you will believe something you regard as less certain than something you don't believe and that is a very odd state of affairs.[3]

Harman and Stalnaker might bite the bullet and say that I need not disbelieve that Clinton will win by exactly twenty votes (or whatever the number may be), despite my holding the negation of this possibility to be more probable than Clinton's winning. After all, they are already committed to the view that you may take the probability of a proposition to be as high as you want, short of unity, without believing it. This means that there could be propositions that I don't believe whose probability for me is greater than that of most of my ordinary beliefs, unless I assign a probability of one to most of my ordinary beliefs (which I don't think I do.)

Considerable violence has to be done to our intuitions to preserve the position staked out by Harman and Stalnaker. Consider, for example, a perfectly ordinary situation such as a glass of lemonade sitting on a picnic table on a warm summer's day. I believe that the lemonade will not be frozen one minute from now, but there are physically possible quantum mechanical states in which it is. The set of quantum mechanical states in which the lemonade is not frozen can be finitely partitioned so that each partition element is of lesser probability than the collection of all states in which the lemonade is frozen one minute from now. If each partition element is less believable than the proposition that the lemonade will be frozen in one minute, then since I disbelieve the latter, I must disbelieve each of the former. But this is impossible by Deductive Closure or the Conjunction Principle because the partition elements, together with the collection of states in which the lemonade is frozen one minute from now, exhaust all the possibilities.

In all these examples, the closest Harman and Stalnaker can come to capturing the epistemic distrust we have in each individual member of the partition, is to say that while one doesn't actually *disbelieve* any partition element, one does believe of each element that it is highly *improbable*. This response does not seem available to them, however, in the case of the famous example of the preface. In this "paradox," we are asked to imagine an author who writes a book in which she makes many claims which she sincerely and fully believes, yet who says in the preface to the book that there are undoubtedly mistakes somewhere in the book. The conjunction of all the propositions expressed in the book, together with the proposition that those are all the propositions expressed in the book, entails that every statement in the book is true. But this is expressly denied by the author in her preface. If the author's beliefs are closed under conjunction and entailment, then she believes an explicit contradiction.

But surely the author does not believe an explicit contradiction (at least not because of what she says in the preface). Why not? Presumably because she fails to believe something entailed by the collection of the beliefs expressed in the book, namely that all these beliefs are true. Hence the Conjunction Principle and the Deductive Closure for beliefs are false. Even if we take these principles to be *normative* principles dictating how people ought to manage their categorical beliefs, it is hard to see that the author has done anything epistemically blameworthy by admitting in the preface that she is fallible. Indeed failure to make such an admission might indicate an unjustified confidence in one's ability to separate truth from falsehood.

Harman and Stalnaker are well aware of the paradox of the preface and the problem it poses for their views. It is much more difficult to respond to this example by denying that the author really believes each statement she makes in the book. How then do they respond? Unfortunately, things get rather murky here. Both Harman and Stalnaker engage in some fancy footwork around the paradox of the preface, but in the end their efforts do not succeed.

Let me begin with Harman's response. I quote in full Harman's remarks about the paradox of the preface:

> We do not ordinarily think of this [the lottery example] as like the case in which an author believes each of the things he or she says in a book he or she has written and also believes that, given human fallibility, at least one of the things he or she has said in the book must be false. Such a person is justified in having inconsistent beliefs, but that does not show that the Recognized Inconsistency Principle is incorrect. It only shows that the principle is defeasible. [pp. 23–24]

In calling the Recognized Inconsistency Principle (what I'm calling "Global Consistency") "defeasible," Harman seems to have something like the following in mind (although he does not spell this out precisely): *normally*, when someone recognizes that certain of her beliefs form an inconsistent set, she has a reason to modify the beliefs in that set. In other words, if someone sees that a particular subset of her beliefs is inconsistent, she has a *prima facie* epistemic obligation to do something to remove the inconsistency. But this obligation is defeasible in the sense that special features of the offending set of beliefs may render the obligation void. So the principle Harman has in mind, I think, is something to the effect that in the normal or typical case in which a set of beliefs is recognized to be inconsistent, there is a reason to modify the beliefs in that set.

Is this principle correct? That's hard to say because it is difficult to say what counts as a "normal" or "typical" case of recognizing that a set of beliefs is inconsistent. Perhaps the normal case is that in which some small number of beliefs, perhaps only two or three, are recognized to be jointly inconsistent. In such cases the principle would hold true, at least approximately, but it would be unclear what guidance it could give us in more complicated cases. Potentially, however, there are lots of cases sufficiently similar to the paradox of the preface in which the principle would fail (e.g. complicated and lengthy mathematical proofs, complex scientific theories, large databases, etc.).

In truth, the interesting question is not whether the principle holds in "normal" cases, but what distinguishes those cases in which it holds from those in which it does not. Harman does not say anything helpful about this latter question. If Harman is correct, there must be some feature of the situation described in the preface paradox that distinguishes it in a relevant way from those situations in which Global Consistency holds. What could that feature be? It could not have anything to do with the content of the beliefs in question, since their content plays no role in the paradox. Surely the relevant feature is not that the beliefs have been committed to writing: most people don't publish their beliefs but still believe, unless they are incredibly cocksure, that not all their beliefs are true. Could the relevant feature be that the belief that causes the inconsistency is a meta-belief, a belief about one's other beliefs? That seems an inessential feature of the situation. There is some proposition P which is the conjunction of all the propositions expressed in the book and although P may be a difficult proposition to hold in one's mind, there is no difficulty in principle in assuming that the author actually believes $\neg P$; if none of the propositions expressed in the book is a proposition about the author's beliefs, then neither will be $\neg P$.

It is true that in the preface example it is more *natural* to think in terms of the meta-belief than in terms of the huge conjunction P, whereas in the lottery example, reference to meta-beliefs seems uncalled for. But I think this merely reflects the availability in the lottery example of a succinct statement equivalent to the negation of a large conjunction. Thus instead of saying "It is not the case that (ticket 1 loses & ticket 2 loses & . . .)," one may simply say "Some ticket will win." In the preface example, no such simple summarizing statement immediately suggests itself. With some thought, though, we can perhaps find simple summarizing statements in the preface example that are not meta-beliefs. Perhaps the author believes that not everything she said happened, did happen, or she may believe that not all the actions attributed to a character in the book were in fact performed by that character. Now these beliefs are not meta-beliefs, but they do involve intentional properties or relations (being said to be the case, being attributed to). Mention of intentional properties, however, seems merely a convenient way of picking out certain events or actions that need not themselves involve anything intentional. Perhaps the author could think referentially about *those* events or *those* actions (i.e. the ones that happen to be described in the book) without thereby thinking about any intentional state.

The only feature of the paradox preface that can plausibly be thought to be responsible for the breakdown of Global Consistency is the beliefs in question being *numerous*. If there are lots of beliefs in question, then a high degree of confidence in each does not imply that one should be confident that all are true. But this is a feature of the lottery example as well: there are a million or more beliefs of the form "Ticket i will not win." In contrast, a five hundred page book will generally not contain more than about ten thousand sentences. So if being numerous is the property of beliefs that explains the failure of Global Consistency in the preface example, it is an even better explanation in the lottery example, where the beliefs are even more numerous.

In contrast to Harman, Stalnaker is not willing to grant that the preface paradox is an exception to any of the principles of belief that he advocates. He defends the Conjunction Principle against the threat posed by the preface paradox as follows (Stalnaker, 1987, p. 94):

> The explanation of the preface phenomenon that I am suggesting requires that we say that the historian does not, without qualification, *believe* [Stalnaker's italics] that the story he accepts is correct; nor does he believe, without qualification, all of the individual statements he makes in telling the story. We must say this to reconcile the phenomenon with a conjunction condition as a rationality condition for belief. But isn't this what we do want to say? The historian, when he wrote his preface, was not just making some additional statement for the reader to believe along with those in his narrative; he was taking something back. It is a conjunction condition for belief, together with a consistency condition, which explains why the reader takes the preface as a hedge or qualification to the text.

Stalnaker is right that the historian does not, without qualification, believe that the story he accepts is correct, if what this means is that he does not believe the story is completely correct. For the historian does not believe that the story is completely correct, and if you don't believe that p, then you don't believe, without qualification, that p (presumably the qualification would be that it might be false that p.) It is harder to understand his claim that the historian does not believe, without qualification, all of the individual statements in his narrative. One way of understanding this is as the claim there is some statement in the story which the historian doesn't really believe but only believes when some proper qualification is added. Now we often do speak of believing something only with a qualification. But typically what this means is that we really don't believe the proposition expressed by a literal, narrow understanding of our words but rather believe some related proposition that is expressed by adding (a sometimes tacit or presupposed) condition to our utterance. If you say "Mary is always punctual," we might say that you don't really believe this without qualification since Mary might not be punctual if there were an unexpected blizzard, or earthquake, or nuclear attack. Thus you really don't believe the proposition that an unqualified reading of your words might suggest, namely the proposition that Mary will be punctual under *all* circumstances. But there's no reason why there must be some statement in the story that the historian believes only with a qualification in this sense. Suppose the historian is very careful to say exactly what he means and always adds the needed qualifications to each statement he makes, so that each statement is something he really and truly does believe without any additional qualification. Will he then believe that by doing so, he has ensured that he made no mistake in the narrative? Of course not.

Perhaps what Stalnaker means is that the statement "The historian believes each statement in the narrative" is true only with a qualification. What might that qualification be? Maybe that although the historian believes each statement, he's not completely sure of each statement. The fact that he realizes that not all the statements in the narrative are true infects each of them with some degree of doubt, we might say. So there must be at least some statements in the story such that the historian is not completely sure that those statements are true.

But even if the historian is not completely sure of some statements in his narrative, he can still believe them in a full-bodied sense of belief. There are very few of my beliefs that I would be willing to stake my life on. In that sense, some degree of doubt infects most of my beliefs. But they're still beliefs. If the Conjunction Principle is restricted to beliefs about which I am absolutely certain, then it is of very limited applicability.

Although neither Harman nor Stalnaker has successfully shown that categorical belief cannot be reduced to probabilistic belief, a problem still remains. The principles of reasoning enunciated by them are exceedingly plausible principles. It is strange to suppose that I believe P and believe Q yet fail to believe P&Q. Yet repeated application of the Conjunction Principle leads to the unacceptable conclusion that one should believe that all of one's beliefs are true. How can such a reasonable principle lead to obviously wrong conclusions? Far from solving this problem by separating categorical belief and probabilistic belief, a probabilistic analysis of belief can explain why these principles are so plausible yet lead to unacceptable conclusions. This is the subject of the next section.

Fuzzy Categorical Belief

Suppose we try to reduce categorical to probabilistic belief by saying that proposition P counts as a belief iff P's subjective probability is at least θ, where θ is a real number less than, but sufficiently close to, one. Call θ

the "threshold" for belief. To give a determinate analysis, we must choose a specific number to be the threshold. The problem is that any fixed threshold for belief seems quite arbitrary. Should 0.95 be the threshold, or 0.99, or something higher? As Stalnaker (1987, p. 91) observes, it seems pointless to single out some specific interval near the top of the probability scale and honor propositions whose probability lies in that interval with the title "believed" while all other propositions are cast into the pit of non-belief. Does it really make sense to suppose that a proposition whose probability is, say, at least 0.98 counts as a belief whereas one whose probability is 0.9799 does not? Surely the difference between these probabilities is insignificant.

I think Stalnaker's observation is correct, but what it shows is that the notion of categorical belief is vague, not that there is no relation between categorical belief and probability. Consider an analogy: There is clearly a relation between the property of being tall and the height of a person. But this relation is not one that can be specified by fixing a precise height beyond which a person is tall and below which he or she is not tall. It does not make sense, for example, to say that someone whose height is at least 1.85 meters is tall whereas someone whose height is only 1.849 meters is not.

Belief is like tallness: it means having a high value of a certain quantity (probability or height) but it is vague exactly what counts as a high value of the quantity. There is no sharp boundary between belief and non-belief; the boundary is instead quite fuzzy. This suggests using some version of fuzzy set theory or of a logic of vagueness to capture the properties of categorical belief. A formal development of the properties of categorical belief using Zadeh's (1975) fuzzy set theory is given in Hunter (1994). In the following, I will keep technical details to a minimum in order not to obscure the general conceptual point.

I claim, then, that categorical belief is simply high subjective probability, where what counts as a "high" probability is fuzzy or vague. So construed, we can ask what principles concerning categorical belief are true. However, if a proposition involves a vague or fuzzy concept, then the question of whether it is true may not have a definite answer. If the color of some object is on the borderline between red and pink, then the statement that it is red is neither clearly true nor clearly false. If the object is somewhat more red than pink, then we might say that it is true to some extent or somewhat true that the object is red. More formally, we may replace the discrete truth-values TRUE and FALSE by a continuous range of truth-values from, say, zero to one, with zero representing definite falsehood, one representing definite truth, and intermediate values representing intermediate degrees of truth. Thus we might assign a degree of truth of 0.5 to the statement "x is red," where the color of x is on the borderline between pink and red. If x's color is somewhat more red than pink but still close to the border between the two colors, then we might assign a degree of truth of 0.7 to the statement "x is red," and so on.[4]

Degrees of truth are not to be confused with probabilities, in *any* sense of "probability." If, for example, we draw randomly from an urn containing equal numbers of black and fiery red balls, the probability of the statement "The next ball drawn will be red" is equal to 0.5, but the degree of truth of that statement will be close to either zero or one. Degrees of truth do not measure the sort of uncertainty dealt with in probability theory, but are more appropriately regarded as a measure of linguistic vagueness.

To model categorical belief as a vague predicate, then, we need some way of assigning degrees of truth to statements of the form "x believes P." How would we do this? I've claimed that to believe a proposition is simply to assign that proposition a high subjective probability, where what counts as a sufficiently high probability is vague. We might try identifying the degree of truth of "x believes P" with x's subjective probability for P, but that seems wrong: a proposition whose subjective probability is 0.5 is not a borderline case of a belief, but is a clear case of a non-belief. A more plausible approach is to take the degree of truth of "x believes P" to be $(2\Pr(P) - 1)^n$, where Pr( ) is x's subjective probability function, and n is an adjustable parameter greater than or equal to one. As n increases, a greater probability for P is required to achieve a given degree of truth for the statement that P is believed; so we can think of n as a parameter determining how high our standards are for what counts as a belief.

Using this definition of vague belief, we can ask about principles concerning categorical belief: How true are they? Is a certain principle definitely true , fairly true, more or less true, or definitely not true?[5] It can be shown that some of the principles proposed by Harman and Stalnaker are always more or less true and often true to a very high degree (Hunter (1994)). Other such principles, though, are sometimes more false than true, but circumstances can be specified in which they too are more true than false.

Let's begin with the principles for belief laid down by Stalnaker. They are:

(S1) If P is believed and P entails Q, then Q is believed.
(S2) If P is believed and Q is believed, then P&Q is believed.
(S3) If P is believed, then not-P is not believed.

(S2) is of course the Conjunction Principle. How true are these principles? (S1) and (S3) are definitely true. (S1) is definitely true since if P entails Q, then Q's probability must be at least as high as P's. (S3) is also definitely true since if P has a high probability, then not-P could not possibly also have a high probability. What about (S2), the troublesome Conjunction

Principle? It's not certainly true, but it's fairly true or approximately true. It says that if P and Q individually have high probabilities, so does P&Q. Is this true? Well, kinda. A lower bound for the probability of P&Q is Pr(P) + Pr(Q) − 1. So if both P and Q have a high probability, say, over 0.9, then the probability of P&Q must be fairly high. For example, if Pr(P) = 0.92 and Pr(Q) = 0.96, then Pr(P&Q) must be at least 0.88 and that's a fairly high probability. Of course, we could have Pr(P) and Pr(Q) each greater than 0.5 while Pr(P&Q) is less than 0.5 (e.g. if Pr(P) = Pr(Q) = 0.7, then we could have Pr(P&Q) = 0.4). But a high probability, at least one sufficient for belief, is not simply one greater than 0.5. A probability of 0.7 isn't really that high, so if both P and Q have a probability greater than 0.5 but not that much greater, then the antecedent of (S2) begins to look false and so that particular instance of (S2) tends to be sort of vacuously true. In general the degree of truth of the consequent tends to vary with the degree of truth of the antecedent in such a way that the entire conditional remains not definitely true, but fairly true (see Hunter (1994)).

So Stalnaker has enunciated three principles for categorical belief, two of which are definitely true and the other of which is close to the truth. That's pretty good. The trouble is that repeated application of a principle that's close to the truth may yield a conclusion that is far from the truth. Thus the famous paradox of the heap proceeds by repeated application of the principle "Take away one stone from a heap and you still have a heap" to derive the clearly false conclusion that zero stones constitute a heap. The trouble is that the "heap" principle is only approximately true and it's approximately, rather than definitely, true due to the vagueness of the concept of a heap. Likewise, repeated application of (S2) leads to obviously false conclusions such as that each of us should believe that all of his or her beliefs are true. Yet this is quite compatible with each individual application of (S2) causing no difficulty.

When we come to the principles advocated by Harman, they turn out not to be even approximately true on a probabilistic analysis of belief. It is not at all close to true that anything entailed by a set of individually highly probable beliefs is itself highly probable. No matter how high a probability you choose, short of one, you can find a set of propositions such that the probability of each proposition in that set exceeds the chosen probability and yet the probability of the conjunction of those propositions is low.[6] Similarly, each member of a set of inconsistent propositions can have a high probability, so Global Consistency is not even approximately true.

How then can we explain the appeal of Deductive Closure and Global Consistency? I think there are several reasons for their intuitive plausibility. First, these principles are consequences of Stalnaker's principles, so if we ignore the difference between being true and being close to true, we may be led to think that Deductive Closure and Global Consistency are just plain true, without qualification.

A second reason has to do with how we in fact make inferences. We generally do not infer a conclusion from a large number of premises in one fell swoop. When there are a large number of premises or initial assumptions, we normally break down the inference into a number of sub-inferences each of which involves only a small number of premises. Formalized systems of deductive reasoning almost always make use of rules of inference with no more than two or three premises. I think it is also true in informal reasoning that a basic inference step involves only a small number of premises. Human beings are not naturally adept at rigorous deductive reasoning anyway, so one would not expect them to make complicated inferences in a single step; nor is it at all natural for human beings to engage in very lengthy chains of deductive reasoning—most reasoning in ordinary contexts is fairly shallow (in a structural sense).[7]

If this is right, then it may explain why Deductive Closure seems plausible. Most cases in which someone recognizes that something follows from other things are cases in which those other things are small in number. In such cases, the conclusion will tend to be highly probable if the premises are, especially if the premises are coherent with one another, as they typically are when someone is trying to present an argument.

Still, there are cases in which people engage in complicated chains of reasoning from large numbers of premises (even if the number of premises in each sub-inference is small). These cases may occur in specialized contexts, but they are certainly not all that uncommon. Why aren't people aware of the breakdown of Deductive Closure and Global Consistency in those sorts of cases? One explanation might be that even when the number of premises is large, the conclusion of the inference may not make full use of all the information in the premises, so that it may be considerably more probable than the conjunction of the premises. Another possible explanation is that people generally adjust upwards their standards for what is believable in order to match the complexity of the problem.

The first reason holds when the conclusion abstracts away from the particular details mentioned in the premises in such a way that it would still hold under changes of those details. To go back to Stalnaker's example of the historian, undoubtedly the historian "is aiming at a coherent total story" (Stalnaker 1987, p. 92) . But the important conclusions that the historian may wish the reader to draw from his narrative are unlikely to be conclusions that depend upon the exact truth of every statement in the narrative. In general, the interesting conclusions will be abstractions from the mass of detail ("The protestant work ethic made capitalism possible," "The severity of the Versailles treaty led to WWII," etc.) and minor changes in those details will make little difference in whether or not the

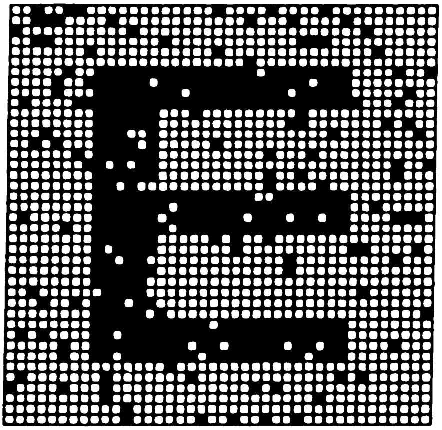

Figure 1. A corrupted letter.

general conclusions are supported. One might say that the historian aims to paint a picture of a certain historical figure or event, a picture that resembles its subject to a very high degree even though it is not a photographic likeness.

The comparison of the historian's narrative to a picture is more than a metaphor. A picture can be thought of as composed of many small bits of colored surface, just as the historian's narrative is composed of very many individual sentences. To borrow some computer science terminology, we might imagine the picture as composed of *pixels*, tiny rectangular elements, each with its own particular color. We form a picture by specifying a color value for each pixel. How we interpret a picture is typically independent of minor changes in the values of pixels. If the picture is of some familiar object, we will still recognize the object even if the values of a few pixels are changed—in other words, a small amount of noise in the picture does not disrupt our ability to interpret the picture. For example, in Figure 1, the capital letter "E" is clearly recognizable despite the surrounding noise. In this case the letter and background are composed of 1600 pixels, approximately ten percent of which have had their color values reversed (from black to white or vice-versa).

I conjecture that human reasoning commonly exhibits the same sort of resiliency in the face of noise even when the reasoning does not concern the recognition of visual patterns. For human inference often has the flavor of "pattern recognition" rather than of straight deductive reasoning.[8] If this is so, then it explains why Deductive Closure *seems* to hold: paradigm cases of human reasoning are cases in which the conclusion is much more probable than the conjunction of the premises. To go back to the case of the historian's narrative, suppose that the important conclusions that the historian wishes to draw from his narrative would still follow if at least, say, ninety percent of what he says in his narrative is true. If each statement in the narrative is probable only to degree 0.95, then it will be almost certain that not all the statements in the narrative are true yet it will be virtually certain that at least ninety percent are true. Here a conclusion may actually be more probable than any individual premise.

Sometimes, however, the truth of all the premises is required in order for the conclusion to be probable. Imagine, for example, a large, complex system such as the space shuttle. Such a system will have many components. In order for the entire system to function properly it is necessary that all or nearly all the components function properly. Thus if one wishes to argue that such a large, complex system will be reliable one must establish that each of its components will behave reliably. It is not sufficient that one establish that a large percentage of the components will be reliable. In this case, if we argue that the system will function properly on the basis of premises stating of each component that it will behave reliably, then our conclusion is no more probable than the conjunction of our premises. Is Deductive Closure at all plausible in a case like this?

I think the answer is that in such a case, where the conclusion depends essentially upon the exact truth of large numbers of assumptions, people tend to be more cautious about accepting each assumption than they would be if only a few assumptions were needed. After all, there's more chance of a mistake if lots of assumptions are involved than if only a few are involved. To ensure that the conclusion is probable, each assumption must be more probable than it would need to be if only a few assumptions were required. Anyone who has done a complicated mathematical proof realizes what great effort is required to ensure that no mistake crops up anywhere, even if each individual step of the proof is a straightforward inference. In contrast, one would rarely devote much time to checking and rechecking a proof with only a few steps. If something similar is true when large numbers of premises are used and their exact truth is crucial, this would explain why people don't notice a failure of Deductive Closure more often—in those cases in which it would fail if the premises were not probable enough, people raise their standards for accepting a premise so that Deductive Closure does not fail.

Something similar may be going on in the case of Global Consistency. What seems to be bad about a set of beliefs being inconsistent is that one might actually put these beliefs together to infer something false or contradictory. But the fact that the set is inconsistent places an upper bound on the probability of each member of the set. This upper bound may not be sufficient to support inferences from large subsets of the set in question.

Finally, we turn to the principle of Rational Monotony. Under what circumstances would Rational Monotony be close to true? It is easy to construct examples in which Rational Monotony is rather more false than true. Recall that Rational Monotony says that if P is believed and $\neg$Q is not believed, then P will still be believed if Q is learned to be true. If we take the statement "P will be believed if Q is learned to be true" to mean that the conditional probability $Pr(P \mid Q)$ is high, then we may take Rational Monotony to claim that when Pr(P) is high and $Pr(\neg Q)$ is not high, then $Pr(P \mid Q)$ is high. This seems fairly false. For example, if we set $Pr(Q) = 0.3$, $Pr(P \mid Q) = 0.7$, and $Pr(P \mid \neg Q) = 1$, then we have $Pr(P) = 0.91$, which seems fairly high, $Pr(\neg Q) = 0.7$ which does not seem high (if there were an almost 1/3 chance of something's happening, we would not normally believe that it would not happen), and yet $Pr(P \mid Q)$ is not high. Unlike the cases of Deductive Closure and Global Consistency, raising standards for belief does not help. If we say that 0.91 is too low for belief and insist that to be believed a proposition's probability must be closer to, say, 0.99 (still leaving things fuzzy—we don't want a precise threshold for belief), then we could set $Pr(Q) = 0.1$, $Pr(P \mid Q) = 0.9$, and $Pr(P \mid \neg Q) = 1$, so that $Pr(P) = 0.99$ and the same conclusion follows—i.e. Rational Monotony is fairly false in this case too.

Part of the problem for Rational Monotony is that to say of a proposition that it is not believed does not imply that its probability does not exceed that of its negation by a significant amount. Presumably a probability of 0.8 is not sufficient for belief yet a proposition with that probability is four times as probable as its negation. It may be that the plausibility of Rational Monotony for some people has stemmed from interpreting "is not believed" to mean something like "has a probability not much greater than 0.5" or "is not more believable than its negation." On such an interpretation, Rational Monotony is close to true.

Suppose, though, we stick with the original formulation of Rational Monotony. Are there circumstances under which it would be approximately true? Yes, there are. If propositions that fail to be believed are in fact propositions whose probabilities are not much greater than 0.5, then Rational Monotony will be approximately true. This would happen if believed or disbelieved propositions tended to have extreme values—either very close to one or close to zero—whereas neither believed nor disbelieved propositions had probabilities close to 0.5. Does this situation obtain?

There is psychological evidence that this situation obtains for many ordinary beliefs. It seems that people have a tendency to be overconfident in their judgments, to be more certain than they should be about the correctness of their beliefs (see Part VI of Kahneman, Slovic, and Tversky 1982). This overconfidence is manifested in, for example, a subject assigning a high degree of confidence (e.g., 90% or more) to each of his answers on a multiple-choice test being correct, while the actual percentage of correct answers is much lower. This overconfidence means that people's subjective probabilities tend to be more extreme than they would be if they had a more realistic view of their own accuracy. This tendency could easily result in the situation described in which probabilities cluster around 0, 0.5, and 1. A proposition for which there is little evidence either for or against might receive a probability around 0.5. As some evidence comes in for or against the proposition, people might overestimate the weight of that evidence and move the probability for the proposition close to either zero or one. In this situation it would be almost as if people were using a crude probability scale with only three values, 0, 0.5, and 1.[9]

The above argument assumes that people are not very good estimators of probabilities. Is there any way we can justify Rational Monotony without making this assumption? In many circumstances, it may be perfectly proper to have extreme subjective probabilities, as, for example, in the case of beliefs about one's immediate surroundings formed by normal perceptual processes or the case of beliefs about one's conscious states. Despite such cases, there are common enough situations in which Rational Monotony fails that some further explanation is required as to why many have found the principle so plausible.

I think the explanation involves an interesting connection between the dynamic principle of Rational Monotony and the static principle of Deductive Closure. If we characterize a belief state simply as a set, namely the set of all propositions believed, then the Deductive Closure principle entails that we could just as well represent the belief state by a single proposition, namely the conjunction of all propositions in the set. Call this proposition "B." By Deductive Closure, B is believed and so is anything entailed by B; furthermore, every belief is entailed by B. Hence a proposition is believed iff it is entailed by B. In this sense, B captures the total content of one's belief state. Suppose now that we adopt a possible worlds representation of propositions: we are given a set of worlds **W** and a proposition is just any subset of **W**. One's belief state B can therefore be represented as a subset of **W**, as shown in Figure 2. Every world in B is compatible with what one believes; every world in ¬B is incompatible with what one believes—i.e. each world in ¬B is disbelieved. Thus the worlds in B have a kind of epistemic priority over worlds in ¬B—the former are at least not disbelieved whereas the latter are all disbelieved. Now suppose one happens to learn some proposition A that is compatible with what one believes now—i.e. A&B, the shaded region in Figure 2, is nonempty. What one believes as a result of learning A will be some subset of A. Which subset? It seems natural to assume that the worlds in A do not change their epistemic priorities as a result of learning A. Hence the worlds in A&B still have epistemic priority over those in A&¬B. So the result of learning A when A is compatible with what one believes will

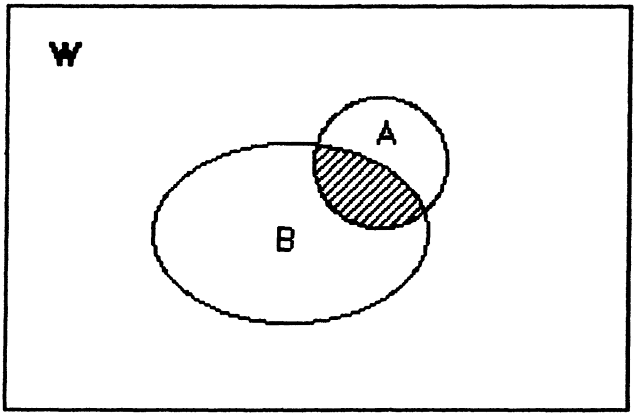

Figure 2. Belief revision in consistent case.

simply be that one's belief state is characterized by the proposition A&B. (Of course, a different rule for revising beliefs would be needed if A is incompatible with what is already believed, since one does not want one's beliefs to be inconsistent.)

If this is how belief revision works when what is learned is compatible with what is already believed, then Rational Monotony follows. For let P be any belief and let Q be a proposition such that ¬Q is not believed. If proposition B characterizes the total belief state, then Q must be compatible with B; otherwise, ¬Q would be entailed by B and so believed. Hence the belief state that results from learning Q is characterized by B&Q. Since proposition P is entailed by B, it is also entailed by B&Q and hence is still believed in the new belief state. That gives us Rational Monotony.

If the argument of this paper is correct, however, this rationale for Rational Monotony collapses. For this rationale assumes that it is appropriate to represent a belief state by a single proposition and this assumption is plausible only if we accept Deductive Closure. Instead of an inviolable principle of belief revision, Rational Monotony turns out to be merely an artifact of a faulty representation of belief.

To sum up the conclusions of this section, we have explored the conditions under which the principles proposed by Harman and Stalnaker, as well as the principle of Rational Monotony, come close to the truth. We have seen that not all these principles are close to the truth under all conditions, but their intuitive plausibility can be explained by the fact that they are very often close to the truth under conditions in which people ordinarily reason with categorical beliefs.

Conclusion

I have argued that categorical belief can be reduced to probabilistic belief and that the plausibility of some principles that apply to categorical belief but seemingly not to probabilistic belief can be explained by these principles being, not definitely true, but true to a high degree. Much more work needs to be done, though, to give a complete account of categorical belief in probabilistic terms. Here I'll mention a few issues that need further discussion.

I've assumed that the agent has a probability function assigning precise probabilities to all propositions of interest. This is psychologically unrealistic: people do not have such precise functions in their heads. Nor do people's degrees of belief always obey the axioms of the probability calculus. I could side-step these issues by confining my analysis to "ideal reasoners," but I prefer to take a different tack. The value of a probabilistic analysis of categorical belief is that it tells us what principles of reasoning are correct, or nearly so, when precise probabilities are lacking, and it also warns us that repeated application of nearly correct principles can lead to trouble. Thus the assumption of an underlying probability function is dispensable: once it has done its job of determining the degree of truth of principles for categorical belief, we can make use of those principles, with appropriate caution, in the absence of precise probabilities.

Another issue that merits more discussion is the context dependence of belief. A proposition that I would be willing to count as a belief in one context may not count as a belief in another, even though my subjective degree of certainty for that belief is the same in both contexts. For example, I may form the opinion that a certain suspect is guilty of murder from reading accounts of the case in the newspaper, but if I were a juror on the case, I might not believe the defendant guilty even given the same evidence. Some might take this to show that belief cannot be defined probabilistically, but I take it to show that different contexts demand different standards for belief. In this respect, belief is similar to many other vague notions. The same individual may be regarded as tall in one context but not tall in another (e.g. when he is a member of a basketball team). This sort of context-dependence of the membership relation for a vague predicate is a familiar issue in fuzzy set theory. The analysis of belief as a vague concept helps us see the context dependence of belief as part of a wider phenomenon.

Finally, I should point out that there is a notion of categorical belief in which the various principles criticized in this paper do hold. This is the notion of belief without any doubt or reservation, which may be identified with a subjective probability equal to, or infinitesimally close to, one. Indeed this seems to be the notion of belief used by Gärdenfors (1988) and Spohn (1988). For this notion of belief, the principles of Deductive Closure, Rational Monotony, et. al. are correct. Other authors, however, apply these principles to belief in the ordinary sense and this is where I object. Ordinary belief should not be identified with probability one (or infinitesimally close to one). Most of my beliefs are not held with absolute certainty, they are not propositions to whose falsity I assign a probability of zero (or an infinitesimally small probability). If we want to understand belief in its workaday sense, rather than in some refined philosophical sense, we should not identify belief with probability equal to, or infinitesimally close to, one.

Notes

*I wish to thank Steve Kemp, Joe Lambert, Reed Richter, and Wolfgang Spohn for helpful comments on an earlier draft of this paper. Thanks also to the anonymous referees for their very constructive comments.

[1]If we assume compactness (that a set of propositions entails P iff some finite subset entails P), then Kyburg's (1970) distinction between the principle of strong consistency and the principle of weak consistency is equivalent to the distinction between global and local consistency.

[2]A somewhat similar example, involving a horse-race, is given in Lehrer (1980). Lehrer, however, firmly holds to both Deductive Closure and Global Consistency and so does not use such examples to argue for a probabilistic rule of acceptance.

An argument against the Conjunction Principle very similar in nature to the one in the present paper was brought to my attention by one of the referees. Korb (1992) derives an even more counterintuitive consequence of that principle than the one I derive from it. In essence, Korb argues that for any contingent proposition P, one can find a finite set **S** of equally probable and mutually exclusive and exhaustive propositions such that P is equivalent to the disjunction of the members in some proper subset **T** of **S**. By the Conjunction Principle, we cannot believe the negation of every member of **S**, else we would believe a contradiction. Since the propositions are equally probable and so in the same epistemic boat, we cannot disbelieve *any* member of **S**. In particular, we cannot disbelieve Q, where $Q \in \mathbf{S}—\mathbf{T}$. But P entails $\neg Q$, so if P is believed, then Q must be disbelieved. Hence P cannot be believed. But P was *any* contingent proposition, so the shocking conclusion is that no contingent proposition is believable if the Conjunction Principle holds true. In a sense, my argument is the same *reductio* of the Conjunction Principle, except that I start with the premise that it is rational to believe at least some contingent propositions. There is, however, an important difference between our arguments. It is essential to Korb's argument that he assume that if two propositions have equal probability, then it is rational to believe one iff it is rational to believe the other. My argument does not assume that the partition elements are equiprobable or that equiprobable propositions must have the same doxastic status; it only assumes that a believed but not completely certain proposition P can be partitioned finely enough so that each partition element is less believable than $\neg P$. This is a weaker and less controversial assumption.

[3]Some might be tempted to think that this argument can be challenged by appealing to Levi's (1967) idea that a particular problem context determines a suitable partition of possibilities to be considered and that different partitions are appropriate in different problem contexts. If our epistemic attitudes are assumed to be relative to the partition of possibilities under consideration, then perhaps the difficulty I cite can be avoided. Thus if it is taken for granted that Clinton will win and the question to be answered is "By how many votes will he win?", the relevant partition of possibilities would be {Clinton wins by one vote, Clinton wins by two votes, . . ., Clinton wins by n votes}, for some large n, and if we have no reason to prefer one of these possibilities over any of the others, we should remain completely agnostic as to whether any particular possibility will be realized. If, however, the question is "Will Clinton win by exactly one vote?", we may take the relevant partition of possibilities to be {Clinton wins by exactly one vote, Clinton does not win by exactly one vote}, and given that

the second possibility is much more probable than the first, we may properly accept it and reject the first.

As Kyburg (1970) points out, however, Levi's rule of acceptance allows us to maintain Deductive Closure and Global Consistency within a problem context but only at the expense of their failure *across* contexts. For example, on Levi's view we may affirm P relative to one problem context yet deny P relative to another. But as Lehrer (1980, p. 188) observes, "a rational person does not think in tight compartments, refusing to consider the answer he gave to one question when answering another."

[4]One might worry that the degree of truth of a statement involving a vague concept is itself vague. Thus it seems overly precise to say that the degree of truth of the statement "x is red," where the color of × is close to the border between pink and red, is exactly 0.6. Why not 0.59 or 0.61? One way to deal with this higher-order vagueness is to "fuzzify" the notion of truth, so that predicates such as "is true," "is fairly true," "is fairly false," and so on, are themselves associated with fuzzy sets. By using such fuzzy truth-predicates to describe the degree of truth of vague statements, we avoid the overprecision of assigning specific numerical degrees of truth to such statements. (Of course, one might wonder why the problem hasn't simply been pushed up to the level of truth-predicates.) For our purposes, this level of sophistication is unnecessary, so I will continue to work with precise degrees of truth.

[5]By "definitely true" I mean having a degree of truth of one; similarly "definitely false" means having a degree of truth of zero. We might define "more or less true" as having a degree of truth greater than 0.5, and so on.

[6]Be careful not to read this claim as contrary to the statement: for any finite and consistent set of premises that entail Q, the probability of Q can be made as high as one wants, short of one, by making the probability of each premise sufficiently high. As Adams (1975) shows, this statement is true. Nonetheless, Adams (1975, p. 2) stressed the point that uncertainties in the premises of a valid argument may be passed in a cumulative fashion to the conclusion, so that when there are many premises, the conclusion may be improbable even though each premise is individually highly probable. If, however, there are few premises, then the conclusion will be probable if the uncertainties in the premises are reasonably small. In a sense, then, Adams anticipated our distinction between a principle such as the Conjunction Principle, which involves only a few premises or beliefs, and more global principles such as Deductive Closure. He did not, however, put the distinction in terms of differences in degrees of truth (in the sense of "degrees of truth" used in fuzzy logic), since he did not analyze belief as a fuzzy concept.

[7]There is a lot of psychological evidence that humans have a quite limited ability to engage in correct deductive reasoning. A number of factors probably contribute to this inability, and one of them appears to be the fact that people find it hard to derive a conclusion when multiple inference steps are needed to do so (see Anderson (1980, p. 325)). This would explain why lengthy chains of inference rarely occur in ordinary reasoning.

[8]Some support for this idea can be found in Johnson-Laird's theory of mental models (Johnson-Laird (1988)), which he has applied both the visual perception and to deductive reasoning (although there are significant differences between these two domains). Neural networks also provide a model of a general-purpose reasoning mechanism that is similar to the process of visual pattern recognition in that a small amount of "noise," or uncertainty, in the data does not affect the conclusion reached. (See Rumelhart, et. al., (1986)). How psychologically plausible neural networks are is another story.

[9]Actually, if the set of beliefs is rich enough, it would be impossible for all beliefs to have probabilities close to 0, 0.5, or 1 unless all beliefs were degenerately zero or one. If there are a number of beliefs with probability 0.5, then some of these are bound to be probabilistically independent of one another. If propositions P and Q are independent, then $Pr(P\&Q) = Pr(P)Pr(Q)$ and $Pr(P \vee Q) = Pr(P) + Pr(Q) - Pr(P)Pr(Q)$. If there are independent propositions whose probabilities are near 0.5, then there are propositions (the conjunctions and disjunctions of pairs of such propositions) whose probabilities are near 0.25 and 0.75. This will not upset the argument of the text if we assume that such propositions can be learned to be true only derivatively, by learning their individual components. In that case, we never have to worry about what happens if we learn a proposition whose probability (before being learned true) is significantly greater than 0.5 but not great enough to count as a belief.

References

Adams, Ernest. (1975) *The Logic of Conditionals*, Dordrecht: Reidel.

Anderson, John. (1980) *Cognitive Psychology and its Implications*, San Francisco: W. H. Freeman and Company.

Gärdenfors, Peter. (1988) *Knowledge in Flux*, Cambridge, Mass.: Bradford Books.

Harman, Gilbert. (1986) *Change in View*, Cambridge, Mass.: Bradford Books.

Hunter, Daniel. (1994) "Deductive and Inductive Reasoning with Fuzzy belief," unpublished manuscript.

Johnson-Laird, Philip. (1988) *The Computer and the Mind*, Cambridge, Mass.: Harvard University Press.

Kahneman, D., Slovic, P., and Tversky, A., editors. (1982) *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press.

Korb, Kevin. (1992) "The Collapse of Collective Defeat: Lessons from the Lottery Paradox," *PSA 1992*, Vol. 1, Philosophy of Science Association, 230–236.

Kyburg, Henry. (1961) *Probability and the Logic of Rational Belief*, Middletown, Conn.: Wesleyan University Press.

Kyburg, Henry. (1970) "Conjunctivitis," in Marshall Swain, ed., *Induction, Acceptance, and Rational Belief*, Dordrecht: Reidel, 55–82.

Lehrer, Keith. (1980) "Coherence and the Racehorse Paradox," in P. A. French, T. E. Uehling, Jr., and H. K. Wettstein, eds., *Midwest Studies in Philosophy, Volume V: Studies in Epistemology*, Minneapolis: University of Minnesota Press.

Levi, Issac. (1967) *Gambling with Truth*, New York: Knopf.

Rumelhart, David, McClelland, James, and the PDP Research Group. (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1, Cambridge, Mass.: MIT Press, A Bradford Book.

Spohn, Wolfgang. (1988) "Ordinal conditional functions: a dynamic theory of epistemic states" in *Causation in Decision, Belief Change, and Statistics*, ed. W. L. Harper and B. Skyrms, Dordrecht: Kluwer, 105–134.

Stalnaker, Robert. (1987) *Inquiry*, Cambridge, Mass.: Bradford Books.

Zadeh, Lofti. (1975) "Fuzzy logic and approximate reasoning," *Synthese* 30 407–428.