First approximations to reasonable models of rationality already exist for categorical inference (Erickson 1978; Johnson-Laird & Steedman 1978; Revlis 1975a) and conditional inference (e.g. Braine 1978; Rips & Marcus 1977; Taplin 1971). These models successfully account for both normatively rational and normatively irrational decisions in experimental settings. In both natural and experimentally constrained settings, decisions appear to be consistent and not to occur at the whim of superficial stimulus characteristics. It is also the case that reasoners' judgments uniformly deviate from the prescriptions of first-order predicate calculus. This is largely the case when the natural quantificational or causal dialect deviates from the normative one. When students are shown sentences that are glossed in the way a logician would gloss them, students' judgments match the logician's (e.g. Adams 1980; Ceraso & Provitera 1971; Revlin, Ammerman, Petersen & Leirer 1978).

Apparent alogicality is either a minor perturbation when problem materials are controlled (Revlin & Leirer 1978) or a direct consequence of the language comprehension processes. While the inferential mechanisms of the head logic may be isomorphic to some normative theory, the encoding component clearly is not. This encoding component reflects primarily the sociology of the experimental setting rather than the character of humans as logical or reasonable processors. Reasoning tasks in the laboratory are properly construed as unilateral dialogues on which the student quite naturally brings normal communicative expectations to bear (e.g. Clark 1977; Grice 1967). It is the logician's dialect in the guise of a psychology experiment that introduces the irrationality into the inferential data.

Cohen notes that untutored reasoners tend to reverse the subject and predicate terms of logical propositions. This conversion of relations is a major contributor to "errors" in formal reasoning (Dickstein 1975; Griggs & Osterman 1980; Revlis 1975b). This encoding is not itself irrational since it is of a piece with rules of normal communication where the symmetrical interpretation is given priority (Bucci 1978; Tsal 1977), whereas circumlocutions are added when asymmetry is intended (Thomas 1976). Conversion of quantified relations occurs in noninferential settings (e.g. Bucci 1978; Revlin & Leirer 1980) and points to the importance of evaluating deductive and inductive decisions in terms of their reasonableness, given the reasoners' inferences about the intentions of messages they are asked to evaluate. The importance of the language-comprehension and encoding components for rational judgments would be obscured had research simply focused on an enumeration of reasoning errors (e.g. Woodworth & Sells 1935) rather than on formal logic problems as a potential dialogue (e.g. Henle 1962).

The fundamental issue then is not whether the appropriate normative model is Baconian or Boolean, but whether or not we proceed to describe decision data in narrow terms – seeking enumeration of "errors," "heuristics," and performance limits or whether we broaden the analysis, examining general principles that may account for human rational competence. It would seem that Cohen's treatment comes down squarely for the latter, which may be in the best interests of the development of reasonably rational models of human inference.

## A theory of probability should tutor our intuitions*

Glenn Shafer

*Department of Mathematics, University of Kansas, Lawrence, Kans. 66045*

L. Jonathan Cohen (1981b) has been a leader in the movement to broaden our perspective on probability. His books *The Implications of Induction* (1970) and *The Probable and the Provable* (1977) were valuable attempts to develop an alternative to the Bayesian theory. But the present paper is unfortunate. The argument for rational competence developed in the first part of the paper is wrong. And the criticisms leveled at Kahneman and Tversky in the second part of the paper are unnecessarily contentious.

Cohen's argument for rational competence leans heavily on an analogy with grammar. A person who has learned a language well can immediately judge whether a sentence is grammatical; one can base a detailed grammar on the evidence provided by these judgments; and it seems reasonable to say that the person has demonstrated his competence to use this grammar, even though his performance may sometimes be faulty. We should deal similarly with probability, Cohen argues. We should build a coherent set of rules or axioms from the evidence provided by ordinary people's probability judgments, and we will then be entitled to say that people have demonstrated their competence to use these rules or axioms.

Unfortunately, the analogy is a poor one. Languages have very rich and – aside from relatively minor ambiguities – definite structures. When a person masters such a structure, it is fair to say that he has mastered a very complex set of rules. The judgments of probability that people make are much less structured and much more confused. And so any set of rules or axioms for probability – any theory of probability – inevitably goes far beyond what could be justified as a mere systematization of "untutored intuitions."

(Cohen puts great stock in untutored intuitions. Are there such things? What a person finds intuitive depends very much on what he has been taught. There is a maxim about this among mathematicians: "Today's intuitions are yesterday's theories.")

Since our intuitions about probability go such a short distance toward determining a theory of probability, there is room to construct many different theories of probability. We have, to cite a few examples, the Bayesian theory, Cohen's Baconian theory, the theory of belief functions (Shafer 1976), and the theory of possibility distributions (Giles 1979; Zadeh 1978). None of these theories can be justified as a mere systematization of lay practice, but they are all theories that a person could learn and then try to use in a constructive way. Each of the theories gives rules relating various probability judgments to one another, and a person could use these rules to construct complex or difficult probability judgments from simpler or easier ones. (See Ekelöf 1981; Shafer 1981).

How, in the end, do we choose among such theories? One criterion is how well and to what advantage people can use the theory. Ultimately, as Cohen argues, we will come back to relatively intuitive judgments; after performing an analysis using the concepts, language, and rules of a particular theory we will have to make a direct judgment as to whether the analysis has given us a better understanding of our problem. But this is not to say that successful theories will owe their success to their congruence with our initial intuitions. On the contrary, they will owe their success to their ability to take us beyond those intuitions.

How well people can use a theory of probability is an empirical question. It will be answered in part by professional users: statisticians, decision analysts, and the like. It will also be answered in part by experimental psychologists.

Cohen's attack on Kahneman and Tversky is part of a quarrel that began in Cohen's *The Probable and the Provable* (1977, pp. 258–64). I find this quarrel unfortunate and unnecessary.

Consider Cohen's attack on Kahneman and Tversky's "heuristics" of availability and representativeness. These heuristics do not have the properties that Cohen's argument for rational competence demands in "human information-processing mechanisms." The argument holds that we are competent in every cognitive task, so that every cognitive task must have an appropriate mechanism. The heuristics are obviously not mechanisms so perfectly fitted to their tasks. Instead, they are tools that can be pressed into service for many tasks, sometimes appropriately, sometimes not. So Cohen attacks heuristics. But is this attack necessary, even from Cohen's own point of view? Why not simply regard the heuristics as tools or subroutines available to more specific mechanisms?

I wish to comment also on Cohen's discussion of Bayes' theorem. As I see it, the fundamental difficulty with the Bayesian analyses Cohen criticizes is that they do not allow for an assessment of the reliability and relevance of "base rates." We are not certain that the predominance of blue cabs in the city as a whole is fully relevant to the particular situation. In a recent paper (Shafer 1982) I discuss how base rates can be discounted using the theory of belief functions; it turns out that one obtains the Bayesian answer if the discount rate is sufficiently low, but that even a moderate discount rate can sharply reduce the influence of the base rate in the face of conflicting evidence. Cohen seems to feel that base rates should always be totally discounted, and this is very difficult to sustain. In Cohen's medical story, for example, the Bayesian analysis seems compelling if there is no reason to discount the relative rarity of disease B.

## Can irrationality be intelligently discussed?

Daniel Kahneman[a] and Amos Tversky[b]

[a]*Department of Psychology, University of British Columbia, Vancouver, B.C., Canada V6T 1W5 and* [b]*Department of Psychology, Stanford University, Stanford, Calif. 94305*

EDITORIAL NOTE

Because Professor Levi's commentary was as much directed to Professors Kahneman and Tversky as to Professor Cohen, the former requested and were given the opportunity to reply at the page-proof stage. Please note that this reply raises some new points directed to Professor Cohen but that it was not possible to give him the chance to see or respond to them within the time constraints of the page-proof stage. Any further discussion will appear in a forthcoming Continuing Commentary section.

We are grateful to *BBS* for the opportunity to make a reply. We note with pleasure that the mood in this round is mellower than it was in the first and that Cohen's concluding remarks are more agreeable than previous installments in our ongoing disagreement of several years' standing. The debate has aired some interesting questions, although we wonder whether the same amount of light could not have been generated with less heat. Evidently, interdisciplinary exchanges are not energy-efficient.

Controversies, especially between participants who specialize in different skills, often resemble informal tennis matches in which the two players keep score independently, with no penalty for hopeful fancy. Such games are safe for the ego. In the case at hand, Cohen finds satisfaction in having compelled us to make what he perceives as a major reformulation of our position, which he now finds less objectionable than the original position that he attributed to us. For his part, he withdraws the guarantee of normative status from lay judgments of probability.

Our score sheet is rather different. In considering the debate as a whole, we are impressed by the paucity of support for Cohen's original position that empirical observations are irrelevant to the analysis of human inductive competence, and by the even weaker support for his attempts to rationalize patterns of judgment such as the neglect of sample size and of base-rate data. On the other hand, we are also impressed by the unease among our colleagues concerning the possibility that our research suggests too bleak a view of the human intellect. We acknowledge that our recent writings reflect a new wariness and a greater effort to avoid misunderstanding, especially by colleagues in other disciplines. The statements to which Cohen refers, however, are not a major reformulation but an attempt to clarify and elaborate an essentially unchanged position that recognizes the complexity of intuitions about chance and uncertainty, the pitfalls of leading questions, and the difference between trivial misunderstandings and serious errors. We note, as Cohen also does, that our positions seem a little closer now than they were earlier. We feel that we have become a little clearer and that his position has become a little more tenable, but, needless to say, we do not expect him to concur.

The current focus of Cohen's disagreement with us is stated on p. 511, where he acknowledges the fallibility of lay judgment, but rejects "intrinsically fallacious mental mechanisms . . . demonstrating some form of cognitive irrationality." These statements attribute to us a position that we never held. Much of our work has been concerned with systematic errors produced by fallible heuristics, but we have never attributed irrationality to our subjects. The discussion of judgmental biases in these loaded terms was in fact initiated by Cohen himself. Furthermore, we have stated many times that heuristics such as availability or representativeness are useful and often yield reasonable estimates. They do not ensure accuracy or coherence but they are neither "intrinsically fallacious" nor "programmed for irrationality."

There now appears to be agreement that the status of judgments as correct or erroneous is in part an empirical question, and that a judgment may be called erroneous if (but not only if) the respondent, after some reflection or discussion, agrees that it was wrong. Because this criterion is difficult to apply in certain problems, we often substitute the shared opinion of experts (e.g., that base rates are relevant) for a direct attempt to educate and test naive subjects. The responses that we have labeled errors satisfy one or both of these tests. The two rounds of commentary indicate that the standards to which we have compared subjects' responses are generally endorsed by the relevant community of experts, with the exception of a few philosophers, all inventors of idiosyncratic and mutually incompatible theories of probability.

The strange debate on the legitimacy of neglecting base-rate data continues. Levi evidently feels that he has laid the issue to rest and Cohen concurs. As might be expected, they are unimpressed by the high frequency of experts who disagree with their position. We agree with Levi that a base rate should not always be used as a prior probability. In the cab problem, for example, a proper Bayesian is free to believe, for whatever reason, that the smaller company causes more than its share of accidents. It seems absurd, however, to go so far as to conclude that "it is a fallacy *not* to neglect the base rate data" (p. 503, italics in the original). Levi's criticism of our analysis of subjects' performance in the cab problem is bewildering. He agrees that standard Bayesian canons require subjects to combine the base rate with the eyewitness report if they recognize the relevance of the base rate in the absence of that report. In fact, we showed that the base rate of cabs is neglected when the eyewitness's testimony is included in the problem whereas it is adopted as a prior probability when the witness is not mentioned. Although we reported the relevant data, Levi does not credit us with the elementary conclusion that the standard Bayesian norms are violated by the conjunction of judgments, rather than by the second judgment alone.

Two main themes were raised in Cohen's target article. The first was that we hold subjects to improper standards. The second was that we misinterpret our subjects' answers to make them appear erroneous. For an example of the strength of the first point, consider yet another base-rate problem.

> A teenage son informed his father that he planned to have his first experience of heroin later that day. The worried father pointed out that statistical research indicates that a majority of teenaged first-time