Psychological Review

VOLUME 90 NUMBER 4 OCTOBER 1983

Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment

Amos Tversky
Stanford University

Daniel Kahneman
University of British Columbia, Vancouver, British Columbia, Canada

Perhaps the simplest and the most basic qualitative law of probability is the conjunction rule: The probability of a conjunction, $P(A\&B)$, cannot exceed the probabilities of its constituents, $P(A)$ and $P(B)$, because the extension (or the possibility set) of the conjunction is included in the extension of its constituents. Judgments under uncertainty, however, are often mediated by intuitive heuristics that are not bound by the conjunction rule. A conjunction can be more representative than one of its constituents, and instances of a specific category can be easier to imagine or to retrieve than instances of a more inclusive category. The representativeness and availability heuristics therefore can make a conjunction appear more probable than one of its constituents. This phenomenon is demonstrated in a variety of contexts including estimation of word frequency, personality judgment, medical prognosis, decision under risk, suspicion of criminal acts, and political forecasting. Systematic violations of the conjunction rule are observed in judgments of lay people and of experts in both between-subjects and within-subjects comparisons. Alternative interpretations of the conjunction fallacy are discussed and attempts to combat it are explored.

Uncertainty is an unavoidable aspect of the human condition. Many significant choices must be based on beliefs about the likelihood of such uncertain events as the guilt of a defendant, the result of an election, the future value of the dollar, the outcome of a medical operation, or the response of a friend. Because we normally do not have adequate formal models for computing the probabilities of such events, intuitive judgment is often the only practical method for assessing uncertainty.

The question of how lay people and experts evaluate the probabilities of uncertain events has attracted considerable research interest in the last decade (see, e.g., Einhorn & Hogarth, 1981; Kahneman, Slovic, & Tversky, 1982; Nisbett & Ross, 1980). Much of this research has compared intuitive inferences and probability judgments to the rules of statistics and the laws of probability. The student of judgment uses the probability calculus as a standard of comparison much as a student of perception might compare the perceived sizes of objects to their physical sizes. Unlike the correct size of objects, however, the "correct" probability of events is not easily defined. Because individuals who have different knowledge or who hold different beliefs must be allowed to assign different probabilities to the same event, no single value can be correct for all people. Furthermore, a correct probability cannot always be determined even for a single person. Outside the domain of random sampling, probability theory does not determine the probabilities of uncertain events—it merely imposes constraints on the relations among

This research was supported by Grant NR197-058 from the U.S. Office of Naval Research. We are grateful to friends and colleagues, too numerous to list by name, for their useful comments and suggestions on an earlier draft of this article.

Requests for reprints should be sent to Amos Tversky, Department of Psychology, Jordan Hall, Building 420, Stanford University, Stanford, California 94305.

them. For example, if A is more probable than B, then the complement of A must be less probable than the complement of B.

The laws of probability derive from extensional considerations. A probability measure is defined on a family of events and each event is construed as a set of possibilities, such as the three ways of getting a 10 on a throw of a pair of dice. The probability of an event equals the sum of the probabilities of its disjoint outcomes. Probability theory has traditionally been used to analyze repetitive chance processes, but the theory has also been applied to essentially unique events where probability is not reducible to the relative frequency of "favorable" outcomes. The probability that the man who sits next to you on the plane is unmarried equals the probability that he is a bachelor plus the probability that he is either divorced or widowed. Additivity applies even when probability does not have a frequentistic interpretation and when the elementary events are not equiprobable.

The simplest and most fundamental qualitative law of probability is the extension rule: If the extension of A includes the extension of B (i.e., $A \supset B$) then $P(A) \geq P(B)$. Because the set of possibilities associated with a conjunction A&B is included in the set of possibilities associated with B, the same principle can also be expressed by the conjunction rule $P(A\&B) \leq P(B)$: A conjunction cannot be more probable than one of its constituents. This rule holds regardless of whether A and B are independent and is valid for any probability assignment on the same sample space. Furthermore, it applies not only to the standard probability calculus but also to nonstandard models such as upper and lower probability (Dempster, 1967; Suppes, 1975), belief function (Shafer, 1976), Baconian probability (Cohen, 1977), rational belief (Kyburg, in press), and possibility theory (Zadeh, 1978).

In contrast to formal theories of belief, intuitive judgments of probability are generally not extensional. People do not normally analyze daily events into exhaustive lists of possibilities or evaluate compound probabilities by aggregating elementary ones. Instead, they commonly use a limited number of heuristics, such as representativeness and availability (Kahneman et al., 1982). Our conception of judgmental heuristics is based on *natural assessments* that are routinely carried out as part of the perception of events and the comprehension of messages. Such natural assessments include computations of similarity and representativeness, attributions of causality, and evaluations of the availability of associations and exemplars. These assessments, we propose, are performed even in the absence of a specific task set, although their results are used to meet task demands as they arise. For example, the mere mention of "horror movies" activates instances of horror movies and evokes an assessment of their availability. Similarly, the statement that Woody Allen's aunt had hoped that he would be a dentist elicits a comparison of the character to the stereotype and an assessment of representativeness. It is presumably the mismatch between Woody Allen's personality and our stereotype of a dentist that makes the thought mildly amusing. Although these assessments are not tied to the estimation of frequency or probability, they are likely to play a dominant role when such judgments are required. The availability of horror movies may be used to answer the question, "What proportion of the movies produced last year were horror movies?", and representativeness may control the judgment that a particular boy is more likely to be an actor than a dentist.

The term *judgmental heuristic* refers to a strategy–whether deliberate or not—that relies on a natural assessment to produce an estimation or a prediction. One of the manifestations of a heuristic is the relative neglect of other considerations. For example, the resemblance of a child to various professional stereotypes may be given too much weight in predicting future vocational choice, at the expense of other pertinent data such as the base-rate frequencies of occupations. Hence, the use of judgmental heuristics gives rise to predictable biases. Natural assessments can affect judgments in other ways, for which the term *heuristic* is less apt. First, people sometimes misinterpret their task and fail to distinguish the required judgment from the natural assessment that the problem evokes. Second, the natural assessment may act as an anchor to which the required judgment is assimiliated, even when the judge does not intend to use the one to estimate the other.

Previous discussions of errors of judgment have focused on deliberate strategies and on misinterpretations of tasks. The present treatment calls special attention to the processes of anchoring and assimiliation, which are often neither deliberate nor conscious. An example from perception may be instructive: If two objects in a picture of a three-dimensional scene have the same picture size, the one that appears more distant is not only seen as "really" larger but also as larger in the picture. The natural computation of real size evidently influences the (less natural) judgment of picture size, although observers are unlikely to confuse the two values or to use the former to estimate the latter.

The natural assessments of representativeness and availability do not conform to the extensional logic of probability theory. In particular, a conjunction can be more representative than one of its constituents, and instances of a specific category can be easier to retrieve than instances of a more inclusive category. The following demonstration illustrates the point. When they were given 60 sec to list seven-letter words of a specified form, students at the University of British Columbia (UBC) produced many more words of the form _ _ _ _ *i n g* than of the form _ _ _ _ _ *n* _, although the latter class includes the former. The average numbers of words produced in the two conditions were 6.4 and 2.9, respectively, $t(44) = 4.70$, $p < .01$. In this test of availability, the increased efficacy of memory search suffices to offset the reduced extension of the target class.

Our treatment of the availability heuristic (Tversky & Kahneman, 1973) suggests that the differential availability of *ing* words and of _ *n* _ words should be reflected in judgments of frequency. The following questions test this prediction.

In four pages of a novel (about 2,000 words), how many words would you expect to find that have the form _ _ _ _ i n g (seven-letter words that end with "ing")? Indicate your best estimate by circling one of the values below:

0 1–2 3–4 5–7 8–10 11–15 16+.

A second version of the question requested estimates for words of the form _ _ _ _ _ *n* _. The median estimates were 13.4 for *ing* words ($n = 52$), and 4.7 for _ *n* _ words ($n = 53$, $p < .01$, by median test), contrary to the extension rule. Similar results were obtained for the comparison of words of the form _ _ _ _ _ *l y* with words of the form _ _ _ _ _ *l* _; the median estimates were 8.8 and 4.4, respectively.

This example illustrates the structure of the studies reported in this article. We constructed problems in which a reduction of extension was associated with an increase in availability or representativeness, and we tested the conjunction rule in judgments of frequency or probability. In the next section we discuss the representativeness heuristic and contrast it with the conjunction rule in the context of person perception. The third section describes conjunction fallacies in medical prognoses, sports forecasting, and choice among bets. In the fourth section we investigate probability judgments for conjunctions of causes and effects and describe conjunction errors in scenarios of future events. Manipulations that enable respondents to resist the conjunction fallacy are explored in the fifth section, and the implications of the results are discussed in the last section.

Representative Conjunctions

Modern research on categorization of objects and events (Mervis & Rosch, 1981; Rosch, 1978; Smith & Medin, 1981) has shown that information is commonly stored and processed in relation to mental models, such as prototypes and schemata. It is therefore natural and economical for the probability of an event to be evaluated by the degree to which that event is representative of an appropriate mental model (Kahneman & Tversky, 1972, 1973; Tversky & Kahneman, 1971, 1982). Because many of the results reported here are attributed to this heuristic, we first briefly analyze the concept of representativeness and illustrate its role in probability judgment.

Representativeness is an assessment of the degree of correspondence between a sample and a population, an instance and a category, an act and an actor or, more generally, between an outcome and a model. The model may refer to a person, a coin, or the world economy, and the respective outcomes could be marital status, a sequence of heads and tails, or the current price of gold. Representativeness can be investigated empirically by asking people, for example, which of two sequences of heads and tails is more representative of a fair coin or which of two professions is more representative of a given personality. This relation differs from other notions of proximity in that it is distinctly directional. It is natural to describe a sample as more or less representative of its parent population or a species (e.g., robin, penguin) as more or less representative of a superordinate category (e.g., bird). It is awkward to describe a population as representative of a sample or a category as representative of an instance.

When the model and the outcomes are described in the same terms, representativeness is reducible to similarity. Because a sample and a population, for example, can be described by the same attributes (e.g., central tendency and variability), the sample appears representative if its salient statistics match the corresponding parameters of the population. In the same manner, a person seems representative of a social group if his or her personality resembles the stereotypical member of that group. Representativeness, however, is not always reducible to similarity; it can also reflect causal and correlational beliefs (see, e.g., Chapman & Chapman, 1967; Jennings, Amabile, & Ross, 1982; Nisbett & Ross, 1980). A particular act (e.g., suicide) is representative of a person because we attribute to the actor a disposition to commit the act, not because the act resembles the person. Thus, an outcome is representative of a model if the salient features match or if the model has a propensity to produce the outcome.

Representativeness tends to covary with frequency: Common instances and frequent events are generally more representative than unusual instances and rare events. The representative summer day is warm and sunny, the representative American family has two children, and the representative height of an adult male is about 5 feet 10 inches. However, there are notable circumstances where representativeness is at variance with both actual and perceived frequency. First, a highly specific outcome can be representative but infrequent. Consider a numerical variable, such as weight, that has a unimodal frequency distribution in a given population. A narrow interval near the mode of the distribution is generally more representative of the population than a wider interval near the tail. For example, 68% of a group of Stanford University undergraduates ($N = 105$) stated that it is more representative for a female Stanford student "to weigh between 124 and 125 pounds" than "to weigh more than 135 pounds". On the other hand, 78% of a different group ($N = 102$) stated that among female Stanford students there are more "women who weigh more than 135 pounds" than "women who weigh between 124 and 125 pounds." Thus, the narrow modal interval (124–125 pounds) was judged to be more representative but less frequent than the broad tail interval (above 135 pounds).

Second, an attribute is representative of a class if it is very diagnostic, that is, if the relative frequency of this attribute is much higher in that class than in a relevant reference class. For example, 65% of the subjects ($N = 105$) stated that it is more representative for a Hollywood actress "to be divorced more than 4 times" than "to vote Democratic." Multiple divorce is diagnostic of Hollywood actresses because it is part of the stereotype that the incidence of divorce is higher among Hollywood actresses than among other women. However, 83% of a different group ($N = 102$) stated that, among Hollywood actresses, there are more "women who vote Democratic" than "women who are divorced more than 4 times." Thus, the more diagnostic attribute was judged to be more representative but less frequent than an attribute (voting Democratic) of lower diagnosticity. Third, an unrepresentative instance of a category can be fairly representative of a superordinate category. For example, chicken is a worse exemplar of a bird than of an animal, and rice is an unrepresentative vegetable, although it is a representative food.

The preceding observations indicate that representativeness is nonextensional: It is not determined by frequency, and it is not bound by class inclusion. Consequently, the test of the conjunction rule in probability judgments offers the sharpest contrast between the extensional logic of probability theory and the psychological principles of representativeness. Our first set of studies of the conjunction rule were conducted in 1974, using occupation and political affiliation as target attributes to be predicted singly or in conjunction from brief personality sketches (see Tversky & Kahneman, 1982, for a brief summary). The studies described in the present section replicate and extend our earlier work. We used the following personality sketches of two fictitious individuals, Bill and Linda, followed by a set of occupations and avocations associated with each of them.

Bill is 34 years old. He is intelligent, but unimaginative, compulsive, and generally lifeless. In school, he was strong in mathematics but weak in social studies and humanities.

Bill is a physician who plays poker for a hobby.
Bill is an architect.
Bill is an accountant. (A)
Bill plays jazz for a hobby. (J)
Bill surfs for a hobby.
Bill is a reporter.
Bill is an accountant who plays jazz for a hobby. (A&J)
Bill climbs mountains for a hobby.

Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Linda is a teacher in elementary school.
Linda works in a bookstore and takes Yoga classes.
Linda is active in the feminist movement. (F)
Linda is a psychiatric social worker.
Linda is a member of the League of Women Voters.
Linda is a bank teller. (T)
Linda is an insurance salesperson.
Linda is a bank teller and is active in the feminist movement. (T&F)

As the reader has probably guessed, the description of Bill was constructed to be representative of an accountant (A) and unrepresentative of a person who plays jazz for a hobby (J). The description of Linda was constructed to be representative of an active feminist (F) and unrepresentative of a bank teller (T). We also expected the ratings of representativeness to be higher for the classes defined by a conjunction of attributes (A&J for Bill, T&F for Linda) than for the less representative constituent of each conjunction (J and T, respectively).

A group of 88 undergraduates at UBC ranked the eight statements associated with each description by "the degree to which Bill (Linda) resembles the typical member of that class." The results confirmed our expectations. The percentages of respondents who displayed the predicted order (A > A&J > J for Bill; F > T&F > T for Linda) were 87% and 85%, respectively. This finding is neither surprising nor objectionable. If, like similarity and prototypicality, representativeness depends on both common and distinctive features (Tversky, 1977), it should be enhanced by the addition of shared features. Adding eyebrows to a schematic face makes it more similar to another schematic face with eyebrows (Gati & Tversky, 1982). Analogously, the addition of feminism to the profession of bank teller improves the match of Linda's current activities to her personality. More surprising and less acceptable is the finding that the great majority of subjects also rank the conjunctions (A&J and T&F) as more *probable* than their less representative constituents (J and T). The following sections describe and analyze this phenomenon.

Indirect and Subtle Tests

Experimental tests of the conjunction rule can be divided into three types: *indirect* tests, *direct-subtle* tests and *direct-transparent* tests. In the indirect tests, one group of subjects evaluates the probability of the conjunction, and another group of subjects evaluates the probability of its constituents. No subject is required to compare a conjunction (e.g., "Linda is a bank teller and a feminist") to its constituents. In the direct-subtle tests, subjects compare the conjunction to its less representative constituent, but the inclusion relation between the events is not emphasized. In the direct-transparent tests, the subjects evaluate or compare the probabilities of the conjunction and its constituent in a format that highlights the relation between them.

The three experimental procedures investigate different hypotheses. The indirect procedure tests whether probability judgments conform to the conjunction rule; the direct-subtle procedure tests whether people will take advantage of an opportunity to compare the critical events; the direct-transparent procedure tests whether people will obey the conjunction rule when they are compelled to compare the critical events. This sequence of tests also describes the course of our investigation, which began with the observation of violations of the conjunction rule in indirect tests and proceeded—to our increasing surprise—to the finding of stubborn failures of that rule in several direct-transparent tests.

Three groups of respondents took part in the main study. The statistically *naive* group consisted of undergraduate students at Stanford University and UBC who had no background in probability or statistics. The *informed* group consisted of first-year graduate students in psychology and in education and of medical students at Stanford who were all familiar with the basic concepts of probability after one or more courses in statistics. The *sophisticated* group consisted of doctoral students in the decision science program of the Stanford Business School who had taken several advanced courses in probability, statistics, and decision theory.

Subjects in the main study received one problem (either Bill or Linda) first in the format of a direct test. They were asked to rank all eight statements associated with that problem (including the conjunction, its separate constituents, and five filler items) according to their probability, using 1 for the most probable and 8 for the least probable. The subjects then received the remaining problem in the format of an indirect test in which the list of alternatives included either the conjunction or its separate constituents. The same five filler items were used in both the direct and the indirect versions of each problem.

Table 1 presents the average ranks (R) of the conjunction R(A&B) and of its less representative constituents R(B), relative to the set of five filler items. The percentage of violations of the conjunction rule in the direct test is denoted by V. The results can be summarized as follows: (a) the conjunction is ranked higher than its less likely constituents in all 12 comparisons, (b) there is no consistent difference between the ranks of the alternatives in the direct and indirect tests, (c) the overall incidence of violations of the conjunction rule in direct tests is 88%, which virtually coincides with the incidence of the corresponding pattern in judgments of representativeness, and (d) there is no effect of statistical sophistication in either indirect or direct tests.

The violation of the conjunction rule in a direct comparison of B to A&B is called the *conjunction fallacy.* Violations inferred from between-subjects comparisons are called *conjunction errors.* Perhaps the most surprising aspect of Table 1 is the lack of any difference between indirect and direct tests. We had expected the conjunction to be judged more probable than the less likely of its constituents in an indirect test, in accord with the pattern observed in judgments of representativeness. However, we also expected that even naive respondents would notice the repetition of some attributes, alone and in conjunction with others, and that they would then apply the conjunction rule and rank the conjunction below its constituents. This expectation was violated, not only by statistically naive undergraduates but even by highly sophisticated respondents. In both direct and indirect tests, the subjects apparently ranked the outcomes by the degree to which Bill (or Linda) matched the respective stereotypes. The correlation between the mean ranks of probability and representativeness was .96 for Bill and .98 for Linda. Does the conjunction rule hold when the relation of inclusion is made highly transparent? The studies described in the next section abandon all subtlety in an effort to compel the subjects to

Table 1
Tests of the Conjunction Rule in Likelihood Rankings

		Direct test				Indirect test		
Subjects	Problem	V	R (A & B)	R (B)	*N*	R (A & B)	R (B)	Total *N*
Naive	Bill	92	2.5	4.5	94	2.3	4.5	88
	Linda	89	3.3	4.4	88	3.3	4.4	86
Informed	Bill	86	2.6	4.5	56	2.4	4.2	56
	Linda	90	3.0	4.3	53	2.9	3.9	55
Sophisticated	Bill	83	2.6	4.7	32	2.5	4.6	32
	Linda	85	3.2	4.3	32	3.1	4.3	32

Note. V = percentage of violations of the conjunction rule; R (A & B) and R (B) = mean rank assigned to A & B and to B, respectively; *N* = number of subjects in the direct test; Total *N* = total number of subjects in the indirect test, who were about equally divided between the two groups.

detect and appreciate the inclusion relation between the target events.

Transparent Tests

This section describes a series of increasingly desperate manipulations designed to induce subjects to obey the conjunction rule. We first presented the description of Linda to a group of 142 undergraduates at UBC and asked them to check which of two alternatives was more probable:

Linda is a bank teller. (T)

Linda is a bank teller and is active in the feminist movement. (T&F)

The order of alternatives was inverted for one half of the subjects, but this manipulation had no effect. Overall, 85% of respondents indicated that T&F was more probable than T, in a flagrant violation of the conjunction rule.

Surprised by the finding, we searched for alternative interpretations of the subjects' responses. Perhaps the subjects found the question too trivial to be taken literally and consequently interpreted the inclusive statement T as T&not-F; that is, "Linda is a bank teller and is *not* a feminist." In such a reading, of course, the observed judgments would not violate the conjunction rule. To test this interpretation, we asked a new group of subjects (N = 119) to assess the probability of T and of T&F on a 9-point scale ranging from 1 (extremely unlikely) to 9 (extremely likely). Because it is sensible to rate probabilities even when one of the events includes the other, there was no reason for respondents to interpret T as T&not-F. The pattern of responses obtained with the new version was the same as before. The mean ratings of probability were 3.5 for T and 5.6 for T&F, and 82% of subjects assigned a higher rating to T&F than they did to T.

Although subjects do not spontaneously apply the conjunction rule, perhaps they can recognize its validity. We presented another group of UBC undergraduates with the description of Linda followed by the two statements, T and T&F, and asked them to indicate which of the following two arguments they found more convincing.

Argument 1. Linda is more likely to be a bank teller than she is to be a feminist bank teller, because every feminist bank teller is a bank teller, but some women bank tellers are not feminists, and Linda could be one of them.

Argument 2. Linda is more likely to be a feminist bank teller than she is likely to be a bank teller, because she resembles an active feminist more than she resembles a bank teller.

The majority of subjects (65%, n = 58) chose the invalid resemblance argument (Argument 2) over the valid extensional argument (Argument 1). Thus, a deliberate attempt to induce a reflective attitude did not eliminate the appeal of the representativeness heuristic.

We made a further effort to clarify the inclusive nature of the event T by representing it as a disjunction. (Note that the conjunction rule can also be expressed as a disjunction rule $P(\text{A or B}) \geq P(\text{B})$). The description of Linda was used again, with a 9-point rating scale for judgments of probability, but the statement T was replaced by

Linda is a bank teller whether or not she is active in the feminist movement. (T*)

This formulation emphasizes the inclusion of T&F in T. Despite the transparent relation between the statements, the mean ratings of likelihood were 5.1 for T&F and 3.8 for T* ($p < .01$, by t test). Furthermore, 57% of the subjects (n = 75) committed the conjunction fallacy by rating T&F higher than T*, and only 16% gave a lower rating to T&F than to T*.

The violations of the conjunction rule in direct comparisons of T&F to T* are remarkable because the extension of "Linda is a bank teller whether or not she is active in the feminist movement" clearly includes the extension of "Linda is a bank teller and is active in the feminist movement." Many subjects evidently failed to draw extensional inferences from the phrase "whether or not," which may have been taken to indicate a weak disposition. This interpretation was supported by a between-subjects comparison, in which different subjects evaluated T, T*, and T&F on a 9-point scale after evaluating the common filler statement, "Linda is a psychiatric social worker." The average ratings were 3.3 for T, 3.9 for T*, and 4.5 for T&F, with each mean significantly different from both others. The statements T and T* are of course extensionally equivalent, but they are assigned different probabilities. Because feminism fits Linda, the mere mention of this attribute makes T* more likely than T, and a definite commitment to it makes the probability of T&F even higher!

Modest success in loosening the grip of the conjunction fallacy was achieved by asking subjects to choose whether to bet on T or on T&F. The subjects were given Linda's description, with the following instruction:

If you could win $10 by betting on an event, which of the following would you choose to bet on? (Check one)

The percentage of violations of the conjunction rule in this task was "only" 56% ($n = 60$), much too high for comfort but substantially lower than the typical value for comparisons of the two events in terms of probability. We conjecture that the betting context draws attention to the conditions in which one bet pays off whereas the other does not, allowing some subjects to discover that a bet on T dominates a bet on T&F.

The respondents in the studies described in this section were statistically naive undergraduates at UBC. Does statistical education eradicate the fallacy? To answer this question, 64 graduate students of social sciences at the University of California, Berkeley and at Stanford University, all with credit for several statistics courses, were given the rating-scale version of the direct test of the conjunction rule for the Linda problem. For the first time in this series of studies, the mean rating for T&F (3.5) was lower than the rating assigned to T (3.8), and only 36% of respondents committed the fallacy. Thus, statistical sophistication produced a majority who conformed to the conjunction rule in a transparent test, although the incidence of violations was fairly high even in this group of intelligent and sophisticated respondents.

Elsewhere (Kahneman & Tversky, 1982a), we distinguished between positive and negative accounts of judgments and preferences that violate normative rules. A positive account focuses on the factors that produce a particular response; a negative account seeks to explain why the correct response was not made. The positive analysis of the Bill and Linda problems invokes the representativeness heuristic. The stubborn persistence of the conjunction fallacy in highly transparent problems, however, lends special interest to the characteristic question of a negative analysis: Why do intelligent and reasonably well-educated people fail to recognize the applicability of the conjunction rule in transparent problems? Postexperimental interviews and class discussions with many subjects shed some light on this question. Naive as well as sophisticated subjects generally noticed the nesting of the target events in the direct-transparent test, but the naive, unlike the sophisticated, did not appreciate its significance for probability assessment. On the other hand, most naive subjects did not attempt to defend their responses. As one subject said after acknowledging the validity of the conjunction rule, "I thought you only asked for my opinion."

The inverviews and the results of the direct transparent tests indicate that naive subjects do not spontaneously treat the conjunction rule as decisive. Their attitude is reminiscent of children's responses in a Piagetian experiment. The child in the preconservation stage is not altogether blind to arguments based on conservation of volume and typically expects quantity to be conserved (Bruner, 1966). What the child fails to see is that the conservation argument is decisive and should overrule the perceptual impression that the tall container holds more water than the short one. Similarly, naive subjects generally endorse the conjunction rule in the abstract, but their application of this rule to the Linda problem is blocked by the compelling impression that T&F is more representative of her than T is. In this context, the adult subjects reason as if they had not reached the stage of formal operations. A full understanding of a principle of physics, logic, or statistics requires knowledge of the conditions under which it prevails over conflicting arguments, such as the height of the liquid in a container or the representativeness of an outcome. The recognition of the decisive nature of rules distinguishes different developmental stages in studies of conservation; it also distinguishes different levels of statistical sophistication in the present series of studies.

More Representative Conjunctions

The preceding studies revealed massive violations of the conjunction rule in the domain of person perception and social stereotypes. Does the conjunction rule fare better in other areas of judgment? Does it hold when the uncertainty regarding the target events is attributed to chance rather than to partial ignorance? Does expertise in the relevant subject matter protect against the conjunction fallacy? Do financial incentives help respondents see the light? The following studies were designed to answer these questions.

Medical Judgment

In this study we asked practicing physicians to make intuitive predictions on the basis of clinical evidence.[1] We chose to study medical judgment because physicians possess expert knowledge and because intuitive judgments often play an important role in medical decision making. Two groups of physicians took part in the study. The first group consisted of 37 internists from the greater Boston area who were taking a postgraduate course at Harvard University. The second group consisted of 66 internists with admitting privileges in the New England Medical Center. They were given problems of the following type:

A 55-year-old woman had pulmonary embolism documented angiographically 10 days after a cholecystectomy.

Please rank order the following in terms of the probability that they will be among the conditions experienced by the patient (use 1 for the most likely and 6 for the least likely). Naturally, the patient could experience more than one of these conditions.

dyspnea and hemiparesis (A&B)	syncope and tachycardia
calf pain	hemiparesis (B)
pleuritic chest pain	hemoptysis

The symptoms listed for each problem included one, denoted B, which was judged by our consulting physicians to be nonrepresentative of the patient's condition, and the conjunction of B with another highly representative symptom denoted A. In the above example of pulmonary embolism (blood clots in the lung), dyspnea (shortness of breath) is a typical symptom, whereas hemiparesis (partial paralysis) is very atypical. Each participant first received three (or two) problems in the indirect format, where the list included either B or the conjunction A&B, but not both, followed by two (or three) problems in the direct format illustrated above. The design was balanced so that each problem appeared about an equal number of times in each format. An independent group of 32 physicians from Stanford University were asked to rank each list of symptoms "by the degree to which they are representative of the clinical condition of the patient."

The design was essentially the same as in the Bill and Linda study. The results of the two experiments were also very similar. The correlation between mean ratings by probability and by representativeness exceeded .95 in all five problems. For every one of the five problems, the conjunction of an unlikely symptom with a likely one was judged more probable than the less likely constituent. The ranking of symptoms was the same in direct and indirect tests: The overall mean ranks of A&B and of B, respectively, were 2.7 and 4.6 in the direct tests and 2.8 and 4.3 in the indirect tests. The incidence of violations of the conjunction rule in direct tests ranged from 73% to 100%, with an average of 91%. Evidently, substantive expertise does not displace representativeness and does not prevent conjunction errors.

Can the results be interpreted without imputing to these experts a consistent violation of the conjunction rule? The instructions used in the present study were especially designed to eliminate the interpretation of Symptom B as an exhaustive description of the relevant facts, which would imply the absence of Symptom A. Participants were instructed to rank symptoms in terms of the probability "that they will be among the conditions experienced by the patient." They were also reminded that "the patient could experience more than one of these conditions." To test the effect of these instructions, the following question was included at the end of the questionnaire:

In assessing the probability that the patient described has a particular symptom X, did you assume that (check one)

X is the *only* symptom experienced by the patient?

X is *among* the symptoms experienced by the patient?

Sixty of the 62 physicians who were asked this question checked the second answer, re-

[1] We are grateful to Barbara J. McNeil, Harvard Medical School, Stephen G. Pauker, Tufts University School of Medicine, and Edward Baer, Stanford Medical School, for their help in the construction of the clinical problems and in the collection of the data.

jecting an interpretation of events that could have justified an apparent violation of the conjunction rule.

An additional group of 24 physicians, mostly residents at Stanford Hospital, participated in a group discussion in which they were confronted with their conjunction fallacies in the same questionnaire. The respondents did not defend their answers, although some references were made to "the nature of clinical experience." Most participants appeared surprised and dismayed to have made an elementary error of reasoning. Because the conjunction fallacy is easy to expose, people who committed it are left with the feeling that they should have known better.

Predicting Wimbledon

The uncertainty encountered in the previous studies regarding the prognosis of a patient or the occupation of a person is normally attributed to incomplete knowledge rather than to the operation of a chance process. Recent studies of inductive reasoning about daily events, conducted by Nisbett, Krantz, Jepson, and Kunda (1983), indicated that statistical principles (e.g., the law of large numbers) are commonly applied in domains such as sports and gambling, which include a random element. The next two studies test the conjunction rule in predictions of the outcomes of a sports event and of a game of chance, where the random aspect of the process is particularly salient.

A group of 93 subjects, recruited through an advertisement in the University of Oregon newspaper, were presented with the following problem in October 1980:

Suppose Bjorn Borg reaches the Wimbledon finals in 1981. Please rank order the following outcomes from most to least likely.

A. Borg will win the match (1.7)
B. Borg will lose the first set (2.7)
C. Borg will lose the first set but win the match (2.2)
D. Borg will win the first set but lose the match (3.5)

The average rank of each outcome (1 = most probable, 2 = second most probable, etc.) is given in parentheses. The outcomes were chosen to represent different levels of strength for the player, Borg, with A indicating the highest strength; C, a rather lower level because it indicates a weakness in the first set; B, lower still because it only mentions this weakness; and D, lowest of all.

After winning his fifth Wimbledon title in 1980, Borg seemed extremely strong. Consequently, we hypothesized that Outcome C would be judged more probable than Outcome B, contrary to the conjunction rule, because C represents a better performance for Borg than does B. The mean rankings indicate that this hypothesis was confirmed; 72% of the respondents assigned a higher rank to C than to B, violating the conjunction rule in a direct test.

Is it possible that the subjects interpreted the target events in a nonextensional manner that could justify or explain the observed ranking? It is well-known that connectives (e.g., *and, or, if*) are often used in ordinary language in ways that depart from their logical definitions. Perhaps the respondents interpreted the conjunction (A and B) as a disjunction (A or B), an implication, (A implies B), or a conditional statement (A if B). Alternatively, the event B could be interpreted in the presence of the conjunction as B and not-A. To investigate these possibilities, we presented to another group of 56 naive subjects at Stanford University the hypothetical results of the relevant tennis match, coded as sequences of wins and losses. For example, the sequence LWWLW denotes a five-set match in which Borg lost (L) the first and the third sets but won (W) the other sets and the match. For each sequence the subjects were asked to examine the four target events of the original Borg problem and to indicate, by marking + or −, whether the given sequence was consistent or inconsistent with each of the events.

With very few exceptions, all of the subjects marked the sequences according to the standard (extensional) interpretation of the target events. A sequence was judged consistent with the conjunction "Borg will lose the first set but win the match" when both constituents were satisfied (e.g., LWWLW) but not when either one or both constituents failed. Evidently, these subjects did not interpret the conjunction as an implication, a conditional statement, or a disjunction. Furthermore, both LWWLW and LWLWL were judged consistent with the inclusive event "Borg will lose the first set," contrary to the hypothesis that the inclusive event B is understood in the context of the other events as "Borg will lose the first set and the match." The classification of sequences therefore indicated little or no ambiguity regarding the extension of the target events. In particular, all sequences that were classified as instances of B&A were also classified as instances of B, but some sequences that were classified as instances of B were judged inconsistent with B&A, in accord with the standard interpretation in which the conjunction rule should be satisfied.

Another possible interpretation of the conjunction error maintains that instead of assessing the probability $P(B/E)$ of Hypothesis B (e.g., that Linda is a bank teller) in light of evidence E (Linda's personality), subjects assess the inverse probability $P(E/B)$ of the evidence given to the hypothesis in question. Because $P(E/A\&B)$ may well exceed $P(E/B)$, the subjects' responses could be justified under this interpretation. Whatever plausibility this account may have in the case of Linda, it is surely inapplicable to the present study where it makes no sense to assess the conditional probability that Borg will reach the finals given the outcome of the final match.

Risky Choice

If the conjunction fallacy cannot be justified by a reinterpretation of the target events, can it be rationalized by a nonstandard conception of probability? On this hypothesis, representativeness is treated as a legitimate nonextensional interpretation of probability rather than as a fallible heuristic. The conjunction fallacy, then, may be viewed as a misunderstanding regarding the meaning of the word *probability.* To investigate this hypothesis we tested the conjunction rule in the following decision problem, which provides an incentive to choose the most probable event, although the word *probability* is not mentioned.

Consider a regular six-sided die with four green faces and two red faces. The die will be rolled 20 times and the sequence of greens (G) and reds (R) will be recorded. You are asked to select one sequence, from a set of three, and you will win $25 if the sequence you chose appears on successive rolls of the die. Please check the sequence of greens and reds on which you prefer to bet.

1. RGRRR
2. GRGRRR
3. GRRRRR

Note that Sequence 1 can be obtained from Sequence 2 by deleting the first G. By the conjunction rule, therefore, Sequence 1 must be more probable than Sequence 2. Note also that all three sequences are rather unrepresentative of the die because they contain more Rs than Gs. However, Sequence 2 appears to be an improvement over Sequence 1 because it contains a higher proportion of the more likely color. A group of 50 respondents were asked to rank the events by the degree to which they are representative of the die; 88% ranked Sequence 2 highest and Sequence 3 lowest. Thus, Sequence 2 is favored by representativeness, although it is dominated by Sequence 1.

A total of 260 students at UBC and Stanford University were given the choice version of the problem. There were no significant differences between the populations, and their results were pooled. The subjects were run in groups of 30 to 50 in a classroom setting. About one half of the subjects (N = 125) actually played the gamble with real payoffs. The choice was hypothetical for the other subjects. The percentages of subjects who chose the dominated option of Sequence 2 were 65% with real payoffs and 62% in the hypothetical format. Only 2% of the subjects in both groups chose Sequence 3.

To facilitate the discovery of the relation between the two critical sequences, we presented a new group of 59 subjects with a (hypothetical) choice problem in which Sequence 2 was replaced by RGRRRG. This new sequence was preferred over Sequence 1, RGRRR, by 63% of the respondents, although the first five elements of the two sequences were identical. These results suggest that subjects coded each sequence in terms of the proportion of Gs and Rs and ranked the sequences by the discrepancy between the proportions in the two sequences (1/5 and 1/3) and the expected value of 2/3.

It is apparent from these results that conjunction errors are not restricted to misunderstandings of the word *probability.* Our subjects followed the representativeness heuristic even when the word was not mentioned and even in choices involving substantial payoffs. The results further show that the conjunction fallacy is not restricted to esoteric interpretations of the connective *and*, because that connective was also absent from the problem. The present test of the conjunction rule was direct, in the sense defined earlier, because the subjects were required to compare two events, one of which included the other. However, informal interviews with some of the respondents suggest that the test was subtle: The relation of inclusion between Sequences 1 and 2 was apparently noted by only a few of the subjects. Evidently, people are not attuned to the detection of nesting among events, even when these relations are clearly displayed.

Suppose that the relation of dominance between Sequences 1 and 2 is called to the subjects' attention. Do they immediately appreciate its force and treat it as a decisive argument for Sequence 1? The original choice problem (without Sequence 3) was presented to a new group of 88 subjects at Stanford University. These subjects, however, were not asked to select the sequence on which they preferred to bet but only to indicate which of the following two arguments, if any, they found correct.

Argument 1: The first sequence (RGRRR) is more probable than the second (GRGRRR) because the second sequence is the same as the first with an additional G at the beginning. Hence, every time the second sequence occurs, the first sequence must also occur. Consequently, you can win on the first and lose on the second, but you can never win on the second and lose on the first.

Argument 2: The second sequence (GRGRRR) is more probable than the first (RGRRR) because the proportions of R and G in the second sequence are closer than those of the first sequence to the expected proportions of R and G for a die with four green and two red faces.

Most of the subjects (76%) chose the valid extensional argument over an argument that formulates the intuition of representativeness. Recall that a similar argument in the case of Linda was much less effective in combating the conjunction fallacy. The success of the present manipulation can be attributed to the combination of a chance setup and a gambling task, which promotes extensional reasoning by emphasizing the conditions under which the bets will pay off.

Fallacies and Misunderstandings

We have described violations of the conjunction rule in direct tests as a fallacy. The term *fallacy* is used here as a psychological hypothesis, not as an evaluative epithet. A judgment is appropriately labeled a fallacy when most of the people who make it are disposed, after suitable explanation, to accept the following propositions: (a) They made a nontrivial error, which they would probably have repeated in similar problems, (b) the error was conceptual, not merely verbal or technical, and (c) they *should* have known the correct answer or a procedure to find it. Alternatively, the same judgment could be described as a failure of communication if the subject misunderstands the question or if the experimenter misinterprets the answer. Subjects who have erred because of a misunderstanding are likely to reject the propositions listed above and to claim (as students often do after an examination) that they knew the correct answer all along, and that their error, if any, was verbal or technical rather than conceptual.

A psychological analysis should apply interpretive charity and should avoid treating genuine misunderstandings as if they were fallacies. It should also avoid the temptation to rationalize any error of judgment by ad hoc interpretations that the respondents themselves would not endorse. The dividing line between fallacies and misunderstandings, however, is not always clear. In one of our earlier studies, for example, most respondents stated that a particular description is more likely to belong to a physical education teacher than to a teacher. Strictly speaking, the latter category includes the former, but it could be argued that *teacher* was understood in this problem in a sense that excludes physical education teacher, much as *animal* is often used in a sense that excludes insects. Hence, it was unclear whether the apparent violation of the extension rule in this problem should be described as a fallacy or as a misunderstanding. A special effort was made in the present studies to avoid ambiguity by defining the critical event as an intersection of well-defined classes, such as bank tellers and feminists. The comments of the respondents in postexperimental discussions supported the conclusion that the observed violations of the conjunction rule in direct tests are genuine fallacies, not just misunderstandings.

Causal Conjunctions

The problems discussed in previous sections included three elements: a causal model M

(Linda's personality); a basic target event B, which is unrepresentative of M (she is a bank teller); and an added event A, which is highly representative of the model M (she is a feminist). In these problems, the model M is positively associated with A and is negatively associated with B. This structure, called the M → A paradigm, is depicted on the left-hand side of Figure 1. We found that when the sketch of Linda's personality was omitted and she was identified merely as a "31-year-old woman," almost all respondents obeyed the conjunction rule and ranked the conjunction (bank teller and active feminist) as less probable than its constituents. The conjunction error in the original problem is therefore attributable to the relation between M and A, not to the relation between A and B.

The conjunction fallacy was common in the Linda problem despite the fact that the stereotypes of bank teller and feminist are mildly incompatible. When the constituents of a conjunction are highly incompatible, the incidence of conjunction errors is greatly reduced. For example, the conjunction "Bill is bored by music and plays jazz for a hobby" was judged as less probable (and less representative) than its constituents, although "bored by music" was perceived as a probable (and representative) attribute of Bill. Quite reasonably, the incompatibility of the two attributes reduced the judged probability of their conjunction.

The effect of compatibility on the evaluation of conjunctions is not limited to near contradictions. For instance, it is more representative (as well as more probable) for a student to be in the upper half of the class in both mathematics and physics or to be in the lower half of the class in both fields than to be in the upper half in one field and in the lower half in the other. Such observations imply that the judged probability (or representativeness) of a conjunction cannot be computed as a function (e.g., product, sum, minimum, weighted average) of the scale values of its constituents. This conclusion excludes a large class of formal models that ignore the relation between the constituents of a conjunction. The viability of such models of conjunctive concepts has generated a spirited debate (Jones, 1982; Osherson & Smith, 1981, 1982; Zadeh, 1982; Lakoff, Note 1).

The preceding discussion suggests a new formal structure, called the A → B paradigm,

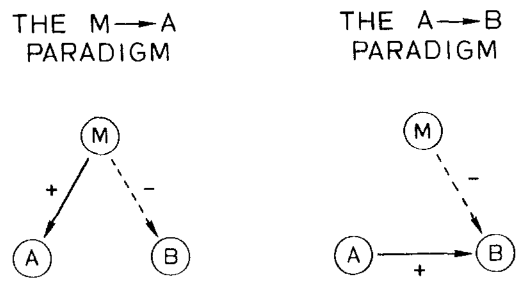

Figure 1. Schematic representation of two experimental paradigms used to test the conjunction rule. (Solid and broken arrows denote strong positive and negative association, respectively, between the model M, the basic target B, and the added target A.)

which is depicted on the right-hand side of Figure 1. Conjunction errors occur in the A → B paradigm because of the direct connection between A and B, although the added event, A, is not particularly representative of the model, M. In this section of the article we investigate problems in which the added event, A, provides a plausible cause or motive for the occurrence of B. Our hypothesis is that the strength of the causal link, which has been shown in previous work to bias judgments of conditional probability (Tversky & Kahneman, 1980), will also bias judgments of the probability of conjunctions (see Beyth-Marom, Note 2). Just as the thought of a personality and a social stereotype naturally evokes an assessment of their similarity, the thought of an effect and a possible cause evokes an assessment of causal impact (Ajzen, 1977). The natural assessment of propensity is expected to bias the evaluation of probability.

To illustrate this bias in the A → B paradigm consider the following problem, which was presented to 115 undergraduates at Stanford University and UBC:

A health survey was conducted in a representative sample of adult males in British Columbia of all ages and occupations.

Mr. F. was included in the sample. He was selected by chance from the list of participants.

Which of the following statements is more probable? (check one)

Mr. F. has had one or more heart attacks.

Mr. F. has had one or more heart attacks and he is over 55 years old.

This seemingly transparent problem elicited a substantial proportion (58%) of conjunction errors among statistically naive respondents. To test the hypothesis that these errors are produced by the causal (or correlational) link between advanced age and heart attacks, rather than by a weighted average of the component probabilities, we removed this link by uncoupling the target events without changing their marginal probabilities.

A health survey was conducted in a representative sample of adult males in British Columbia of all ages and occupations.

Mr. F. and Mr. G. were both included in the sample. They were unrelated and were selected by chance from the list of participants.

Which of the following statements is more probable? (check one)

Mr. F. has had one or more heart attacks.

Mr. F has had one or more heart attacks and Mr. G. is over 55 years old.

Assigning the critical attributes to two independent individuals eliminates in effect the A → B connection by making the events (conditionally) independent. Accordingly, the incidence of conjunction errors dropped to 29% (N = 90).

The A → B paradigm can give rise to dual conjunction errors where A&B is perceived as more probable than each of its constituents, as illustrated in the next problem.

Peter is a junior in college who is training to run the mile in a regional meet. In his best race, earlier this season, Peter ran the mile in 4:06 min. Please rank the following outcomes from most to least probable.

Peter will run the mile under 4:06 min.

Peter will run the mile under 4 min.

Peter will run the second half-mile under 1:55 min.

Peter will run the second half-mile under 1:55 min. and will complete the mile under 4 min.

Peter will run the first half-mile under 2:05 min.

The critical event (a sub-1:55 minute second half *and* a sub-4 minute mile) is clearly defined as a conjunction and not as a conditional. Nevertheless, 76% of a group of undergraduate students from Stanford University (N = 96) ranked it above one of its constituents, and 48% of the subjects ranked it above both constituents. The natural assessment of the relation between the constituents apparently contaminated the evaluation of their conjunction. In contrast, no one violated the extension rule by ranking the second outcome (a sub-4 minute mile) above the first (a sub-4:06 minute mile). The preceding results indicate that the judged probability of a conjunction cannot be explained by an averaging model because in such a model P(A&B) lies between P(A) and P(B). An averaging process, however, may be responsible for some conjunction errors, particularly when the constituent probabilities are given in a numerical form.

Motives and Crimes

A conjunction error in a motive–action schema is illustrated by the following problem—one of several of the same general type administered to a group of 171 students at UBC:

John P. is a meek man, 42 years old, married with two children. His neighbors describe him as mild-mannered, but somewhat secretive. He owns an import–export company based in New York City, and he travels frequently to Europe and the Far East. Mr. P. was convicted once for smuggling precious stones and metals (including uranium) and received a suspended sentence of 6 months in jail and a large fine.

Mr. P. is currently under police investigation.

Please rank the following statements by the probability that they will be among the conclusions of the investigation. Remember that other possibilities exist and that more than one statement may be true. Use 1 for the most probable statement, 2 for the second, etc.

Mr. P. is a child molester.

Mr. P. is involved in espionage and the sale of secret documents.

Mr. P. is a drug addict.

Mr. P. killed one of his employees.

One half of the subjects (n = 86) ranked the events above. Other subjects (n = 85) ranked a modified list of possibilities in which the last event was replaced by

Mr. P. killed one of his employees to prevent him from talking to the police.

Although the addition of a possible motive clearly reduces the extension of the event (Mr. P. might have killed his employee for other reasons, such as revenge or self-defense), we hypothesized that the mention of a plausible but nonobvious motive would increase the perceived likelihood of the event. The data confirmed this expectation. The mean rank of the conjunction was 2.90, whereas the mean rank of the inclusive statement was 3.17 ($p <$ .05, by t test). Furthermore, 50% of the respondents ranked the conjunction as more likely than the event that Mr. P. was a drug addict, but only 23% ranked the more inclusive target event as more likely than drug addiction. We have found in other problems of the same type that the mention of a cause or motive tends to increase the judged probability of an action when the suggested motive (a) offers a reasonable explanation of the target event, (b) appears fairly likely on its own, (c) is nonobvious, in the sense that it does not immediately come to mind when the outcome is mentioned.

We have observed conjunction errors in other judgments involving criminal acts in both the A $\rightarrow$ B and the M $\rightarrow$ A paradigms. For example, the hypothesis that a policeman described as violence prone was involved in the heroin trade was ranked less likely (relative to a standard comparison set) than a conjunction of allegations—that he is involved in the heroin trade and that he recently assaulted a suspect. In that example, the assault was not causally linked to the involvement in drugs, but it made the combined allegation more representative of the suspect's disposition. The implications of the psychology of judgment to the evaluation of legal evidence deserve careful study because the outcomes of many trials depend on the ability of a judge or a jury to make intuitive judgments on the basis of partial and fallible data (see Rubinstein, 1979; Saks & Kidd, 1981).

Forecasts and Scenarios

The construction and evaluation of scenarios of future events are not only a favorite pastime of reporters, analysts, and news watchers. Scenarios are often used in the context of planning, and their plausibility influences significant decisions. Scenarios for the past are also important in many contexts, including criminal law and the writing of history. It is of interest, then, to evaluate whether the forecasting or reconstruction of real-life events is subject to conjunction errors. Our analysis suggests that a scenario that includes a possible cause and an outcome could appear more probable than the outcome on its own. We tested this hypothesis in two populations: statistically naive students and professional forecasters.

A sample of 245 UBC undergraduates were requested in April 1982 to evaluate the probability of occurrence of several events in 1983. A 9-point scale was used, defined by the following categories: less than .01%, .1%, .5%, 1%, 2%, 5%, 10%, 25%, and 50% or more. Each problem was presented to different subjects in two versions: one that included only the basic outcome and another that included a more detailed scenario leading to the same outcome. For example, one half of the subjects evaluated the probability of

> a massive flood somewhere in North America in 1983, in which more than 1000 people drown.

The other half of the subjects evaluated the probability of

> an earthquake in California sometime in 1983, causing a flood in which more than 1000 people drown.

The estimates of the conjunction (earthquake and flood) were significantly higher than the estimates of the flood ($p < .01$, by a Mann–Whitney test). The respective geometric means were 3.1% and 2.2%. Thus, a reminder that a devastating flood could be caused by the anticipated California earthquake made the conjunction of an earthquake and a flood appear more probable than a flood. The same pattern was observed in other problems.

The subjects in the second part of the study were 115 participants in the Second International Congress on Forecasting held in Istanbul, Turkey in July 1982. Most of the subjects were professional analysts, employed by industry, universities, or research institutes. They were professionally involved in forecasting and planning, and many had used scenarios in their work. The research design and the response scales were the same as before. One group of forecasters evaluated the probability of

> a complete suspension of diplomatic relations between the USA and the Soviet Union, sometime in 1983.

The other respondents evaluated the probability of the same outcome embedded in the following scenario:

> a Russian invasion of Poland, and a complete suspension of diplomatic relations between the USA and the Soviet Union, sometime in 1983.

Although *suspension* is necessarily more probable than *invasion and suspension,* a Russian invasion of Poland offered a plausible scenario leading to the breakdown of diplomatic relations between the superpowers. As expected, the estimates of probability were low for both problems but significantly higher for the conjunction *invasion and suspension* than for *suspension* ($p < .01$, by a Mann–Whitney test). The geometric means of estimates were .47% and .14%, respectively. A similar effect was observed in the comparison of the following outcomes:

a 30% drop in the consumption of oil in the US in 1983.

a dramatic increase in oil prices and a 30% drop in the consumption of oil in the US in 1983.

The geometric means of the estimated probability of the first and the second outcomes, respectively, were .22% and .36%. We speculate that the effect is smaller in this problem (although still statistically significant) because the basic target event (a large drop in oil consumption) makes the added event (a dramatic increase in oil prices) highly available, even when the latter is not mentioned.

Conjunctions involving hypothetical causes are particularly prone to error because it is more natural to assess the probability of the effect given the cause than the joint probability of the effect and the cause. We do not suggest that subjects deliberately adopt this interpretation; rather we propose that the higher conditional estimate serves as an anchor that makes the conjunction appear more probable.

Attempts to forecast events such as a major nuclear accident in the United States or an Islamic revolution in Saudi Arabia typically involve the construction and evaluation of scenarios. Similarly, a plausible story of how the victim might have been killed by someone other than the defendant may convince a jury of the existence of reasonable doubt. Scenarios can usefully serve to stimulate the imagination, to establish the feasibility of outcomes, or to set bounds on judged probabilities (Kirkwood & Pollock, 1982; Zentner, 1982). However, the use of scenarios as a prime instrument for the assessment of probabilities can be highly misleading. First, this procedure favors a conjunctive outcome produced by a sequence of likely steps (e.g., the successful execution of a plan) over an equally probable disjunctive outcome (e.g., the failure of a careful plan), which can occur in many unlikely ways (Bar-Hillel, 1973; Tversky & Kahneman, 1973). Second, the use of scenarios to assess probability is especially vulnerable to conjunction errors. A detailed scenario consisting of causally linked and representative events may appear more probable than a subset of these events (Slovic, Fischhoff, & Lichtenstein, 1976). This effect contributes to the appeal of scenarios and to the illusory insight that they often provide. The attorney who fills in guesses regarding unknown facts, such as motive or mode of operation, may strengthen a case by improving its coherence, although such additions can only lower probability. Similarly, a political analyst can improve scenarios by adding plausible causes and representative consequences. As Pooh-Bah in the *Mikado* explains, such additions provide "corroborative details intended to give artistic verisimilitude to an otherwise bald and unconvincing narrative."

Extensional Cues

The numerous conjunction errors reported in this article illustrate people's affinity for nonextensional reasoning. It is nonetheless obvious that people can understand and apply the extension rule. What cues elicit extensional considerations and what factors promote conformity to the conjunction rule? In this section we focus on a single estimation problem and report several manipulations that induce extensional reasoning and reduce the incidence of the conjunction fallacy. The participants in the studies described in this section were statistically naive students at UBC. Mean estimates are given in parentheses.

A health survey was conducted in a sample of adult males in British Columbia, of all ages and occupations.

Please give your best estimate of the following values:

What percentage of the men surveyed have had one or more heart attacks? (18%)

What percentage of the men surveyed both are over 55 years old and have had one or more heart attacks? (30%)

This version of the health-survey problem produced a substantial number of conjunction errors among statistically naive respondents: 65% of the respondents ($N = 147$) assigned a strictly higher estimate to the second question than to the first.[2] Reversing the order of the constituents did not significantly affect the results.

The observed violations of the conjunction rule in estimates of relative frequency are attributed to the A → B paradigm. We propose that the probability of the conjunction is biased toward the natural assessment of the strength of the causal or statistical link between age and heart attacks. Although the statement of the question appears unambiguous, we considered the hypothesis that the respondents who committed the fallacy had actually interpreted the second question as a request to assess a conditional probability. A new group of UBC undergraduates received the same problem, with the second question amended as follows:

Among the men surveyed who are over 55 years old, what percentage have had one or more heart attacks?

The mean estimate was 59% (N = 55). This value is significantly higher than the mean of the estimates of the conjunction (45%) given by those subjects who had committed the fallacy in the original problem. Subjects who violate the conjunction rule therefore do not simply substitute the conditional P(B/A) for the conjunction P(A&B).

A seemingly inconsequential change in the problem helps many respondents avoid the conjunction fallacy. A new group of subjects (N = 159) were given the original questions but were also asked to assess the "percentage of the men surveyed who are over 55 years old" prior to assessing the conjunction. This manipulation reduced the incidence of conjunction error from 65% to 31%. It appears that many subjects were appropriately cued by the requirement to assess the relative frequency of both classes before assessing the relative frequency of their intersection.

The following formulation also facilitates extensional reasoning:

A health survey was conducted in a sample of 100 adult males in British Columbia, of all ages and occupations.

Please give your best estimate of the following values:

How many of the 100 participants have had one or more heart attacks?

How many of the 100 participants both are over 55 years old and have had one or more heart attacks?

The incidence of the conjunction fallacy was only 25% in this version (N = 117). Evidently, an explicit reference to the number of individual cases encourages subjects to set up a representation of the problems in which class inclusion is readily perceived and appreciated. We have replicated this effect in several other problems of the same general type. The rate of errors was further reduced to a record 11% for a group (N = 360) who also estimated the number of participants over 55 years of age prior to the estimation of the conjunctive category. The present findings agree with the results of Beyth-Marom (Note 2), who observed higher estimates for conjunctions in judgments of probability than in assessments of frequency.

The results of this section show that nonextensional reasoning sometimes prevails even in simple estimates of relative frequency in which the extension of the target event and the meaning of the scale are completely unambiguous. On the other hand, we found that the replacement of percentages by frequencies and the request to assess both constituent categories markedly reduced the incidence of the conjunction fallacy. It appears that extensional considerations are readily brought to mind by seemingly inconsequential cues. A contrast worthy of note exists between the effectiveness of extensional cues in the health-survey problem and the relative inefficacy of the methods used to combat the conjunction fallacy in the Linda problem (argument, betting, "whether or not"). The force of the conjunction rule is more readily appreciated when the conjunctions are defined by the intersection of concrete classes than by a combination of properties. Although classes and properties are equivalent from a logical standpoint, they give rise to different mental representations in which different relations and rules are transparent. The formal equivalence of properties to classes is apparently not programmed into the lay mind.

Discussion

In the course of this project we studied the extension rule in a variety of domains; we tested more than 3,000 subjects on dozens of

[2] The incidence of the conjunction fallacy was considerably lower (28%) for a group of advanced undergraduates at Stanford University (N = 62) who had completed one or more courses in statistics.

problems, and we examined numerous variations of these problems. The results reported in this article constitute a representative though not exhaustive summary of this work.

The data revealed widespread violations of the extension rule by naive and sophisticated subjects in both indirect and direct tests. These results were interpreted within the framework of judgmental heuristics. We proposed that a judgment of probability or frequency is commonly biased toward the natural assessment that the problem evokes. Thus, the request to estimate the frequency of a class elicits a search for exemplars, the task of predicting vocational choice from a personality sketch evokes a comparison of features, and a question about the co-occurrence of events induces an assessment of their causal connection. These assessments are not constrained by the extension rule. Although an arbitrary reduction in the extension of an event typically reduces its availability, representativeness, or causal coherence, there are numerous occasions in which these assessments are higher for the restricted than for the inclusive event. Natural assessments can bias probability judgment in three ways: The respondents (a) may use a natural assessment deliberately as a strategy of estimation, (b) may be primed or anchored by it, or (c) may fail to appreciate the difference between the natural and the required assessments.

Logic Versus Intuition

The conjunction error demonstrates with exceptional clarity the contrast between the extensional logic that underlies most formal conceptions of probability and the natural assessments that govern many judgments and beliefs. However, probability judgments are not always dominated by nonextensional heuristics. Rudiments of probability theory have become part of the culture, and even statistically naive adults can enumerate possibilities and calculate odds in simple games of chance (Edwards, 1975). Furthermore, some real-life contexts encourage the decomposition of events. The chances of a team to reach the playoffs, for example, may be evaluated as follows: "Our team will make it if we beat team B, which we should be able to do since we have a better defense, or if team B loses to both C and D, which is unlikely since neither one has a strong offense." In this example, the target event (reaching the playoffs) is decomposed into more elementary possibilities that are evaluated in an intuitive manner.

Judgments of probability vary in the degree to which they follow a decompositional or a holistic approach and in the degree to which the assessment and the aggregation of probabilities are analytic or intuitive (see, e.g., Hammond & Brehmer, 1973). At one extreme there are questions (e.g., What are the chances of beating a given hand in poker?) that can be answered by calculating the relative frequency of "favorable" outcomes. Such an analysis possesses all the features associated with an extensional approach: It is decompositional, frequentistic, and algorithmic. At the other extreme, there are questions (e.g., What is the probability that the witness is telling the truth?) that are normally evaluated in a holistic, singular, and intuitive manner (Kahneman & Tversky, 1982b). Decomposition and calculation provide some protection against conjunction errors and other biases, but the intuitive element cannot be entirely eliminated from probability judgments outside the domain of random sampling.

A direct test of the conjunction rule pits an intuitive impression against a basic law of probability. The outcome of the conflict is determined by the nature of the evidence, the formulation of the question, the transparency of the event structure, the appeal of the heuristic, and the sophistication of the respondents. Whether people obey the conjunction rule in any particular direct test depends on the balance of these factors. For example, we found it difficult to induce naive subjects to apply the conjunction rule in the Linda problem, but minor variations in the health-survey question had a marked effect on conjunction errors. This conclusion is consistent with the results of Nisbett et al. (1983), who showed that lay people can apply certain statistical principles (e.g., the law of large numbers) to everyday problems and that the accessibility of these principles varied with the content of the problem and increased significantly with the sophistication of the respondents. We found, however, that sophisticated and naive respondents answered the Linda problem similarly in indirect tests and only parted company in the most transparent versions of the problem. These observations suggest that statistical sophistication did not alter intuitions of representativeness, although it enabled the respondents to recognize in direct tests the decisive force of the extension rule.

Judgment problems in real life do not usually present themselves in the format of a within-subjects design or of a direct test of the laws of probability. Consequently, subjects' performance in a between-subjects test may offer a more realistic view of everyday reasoning. In the indirect test it is very difficult even for a sophisticated judge to ensure that an event has no subset that would appear more probable than it does and no superset that would appear less probable. The satisfaction of the extension rule could be ensured, without direct comparisons of A&B to B, if all events in the relevant ensemble were expressed as disjoint unions of elementary possibilities. In many practical contexts, however, such analysis is not feasible. The physician, judge, political analyst, or entrepreneur typically focuses on a critical target event and is rarely prompted to discover potential violations of the extension rule.

Studies of reasoning and problem solving have shown that people often fail to understand or apply an abstract logical principle even when they can use it properly in concrete familiar contexts. Johnson-Laird and Wason (1977), for example, showed that people who err in the verification of *if then* statements in an abstract format often succeed when the problem evokes a familiar schema. The present results exhibit the opposite pattern: People generally accept the conjunction rule in its abstract form (B is more probable than A&B) but defy it in concrete examples, such as the Linda and Bill problems, where the rule conflicts with an intuitive impression.

The violations of the conjunction rule were not only prevalent in our research, they were also sizable. For example, subjects' estimates of the frequency of seven-letter words ending with *ing* were three times as high as their estimates of the frequency of seven letter words ending with _ *n* _. A correction by a factor of three is the smallest change that would eliminate the inconsistency between the two estimates. However, the subjects surely know that there are many _ *n* _ words that are not *ing* words (e.g., *present, content*). If they believe, for example, that only one half of the _ *n* _ words end with *ing*, then a 6:1 adjustment would be required to make the entire system coherent. The ordinal nature of most of our experiments did not permit an estimate of the adjustment factor required for coherence. Nevertheless, the size of the effect was often considerable. In the rating-scale version of the Linda problem, for example, there was little overlap between the distributions of ratings for T&F and for T. Our problems, of course, were constructed to elicit conjunction errors, and they do not provide an unbiased estimate of the prevalence of these errors. Note, however, that the conjunction error is only a symptom of a more general phenomenon: People tend to overestimate the probabilities of representative (or available) events and/or underestimate the probabilities of less representative events. The violation of the conjunction rule demonstrates this tendency even when the "true" probabilities are unknown or unknowable. The basic phenomenon may be considerably more common than the extreme symptom by which it was illustrated.

Previous studies of the subjective probability of conjunctions (e.g., Bar-Hillel, 1973; Cohen & Hansel, 1957; Goldsmith, 1978; Wyer, 1976; Beyth-Marom, Note 2) focused primarily on testing the multiplicative rule $P(\text{A\&B}) = P(\text{B})P(\text{A/B})$. This rule is strictly stronger than the conjunction rule; it also requires cardinal rather than ordinal assessments of probability. The results showed that people generally overestimate the probability of conjunctions in the sense that $P(\text{A\&B}) > P(\text{B})P(\text{A/B})$. Some investigators, notably Wyer and Beyth-Marom, also reported data that are inconsistent with the conjunction rule.

Conversing Under Uncertainty

The representativeness heuristic generally favors outcomes that make good stories or good hypotheses. The conjunction *feminist bank teller* is a better hypothesis about Linda than *bank teller*, and the scenario of a Russian invasion of Poland followed by a diplomatic crisis makes a better story than simply *diplomatic crisis*. The notion of a good story can be illuminated by extending the Gricean concept of cooperativeness (Grice, 1975) to conversations under uncertainty. The standard analysis of conversation rules assumes that the speaker knows the truth. The maxim of quality enjoins him or her to say only the truth. The maxim of quantity enjoins the speaker to say all of it, subject to the maxim of relevance, which restricts the message to what the listener needs to know. What rules of cooperativeness apply to an uncertain speaker, that is, one who is uncertain of the truth? Such a speaker can guarantee absolute quality only for tautological statements (e.g., "Inflation will continue so long as prices rise"), which are unlikely to earn high marks as contributions to the conversation. A useful contribution must convey the speaker's relevant beliefs even if they are not certain. The rules of cooperativeness for an uncertain speaker must therefore allow for a trade-off of quality and quantity in the evaluation of messages. The expected value of a message can be defined by its information value if it is true, weighted by the probability that it is true. An uncertain speaker may wish to follow the maxim of value: Select the message that has the highest expected value.

The expected value of a message can sometimes be improved by increasing its content, although its probability is thereby reduced. The statement "Inflation will be in the range of 6% to 9% by the end of the year" may be a more valuable forecast than "Inflation will be in the range of 3% to 12%," although the latter is more likely to be confirmed. A good forecast is a compromise between a point estimate, which is sure to be wrong, and a 99.9% credible interval, which is often too broad. The selection of hypotheses in science is subject to the same trade-off: A hypothesis must risk refutation to be valuable, but its value declines if refutation is nearly certain. Good hypotheses balance informativeness against probable truth (Good, 1971). A similar compromise obtains in the structure of natural categories. The basic level category *dog* is much more informative than the more inclusive category *animal* and only slightly less informative than the narrower category *beagle.* Basic level categories have a privileged position in language and thought, presumably because they offer an optimal combination of scope and content (Rosch, 1978). Categorization under uncertainty is a case in point. A moving object dimly seen in the dark may be appropriately labeled *dog,* where the subordinate *beagle* would be rash and the superordinate *animal* far too conservative.

Consider the task of ranking possible answers to the question, "What do you think Linda is up to these days?" The maxim of value could justify a preference for T&F over T in this task, because the added attribute *feminist* considerably enriches the description of Linda's current activities, at an acceptable cost in probable truth. Thus, the analysis of conversation under uncertainty identifies a pertinent question that is legitimately answered by ranking the conjunction above its constituent. We do not believe, however, that the maxim of value provides a fully satisfactory account of the conjunction fallacy. First, it is unlikely that our respondents interpret the request to rank statements by their probability as a request to rank them by their expected (informational) value. Second, conjunction fallacies have been observed in numerical estimates and in choices of bets, to which the conversational analysis simply does not apply. Nevertheless, the preference for statements of high expected (informational) value could hinder the appreciation of the extension rule. As we suggested in the discussion of the interaction of picture size and real size, the answer to a question can be biased by the availability of an answer to a cognate question—even when the respondent is well aware of the distinction between them.

The same analysis applies to other conceptual neighbors of probability. The concept of surprise is a case in point. Although surprise is closely tied to expectations, it does not follow the laws of probability (Kahneman & Tversky, 1982b). For example, the message that a tennis champion lost the first set of a match is more surprising than the message that she lost the first set but won the match, and a sequence of four consecutive heads in a coin toss is more surprising than four heads followed by two tails. It would be patently absurd, however, to bet on the less surprising event in each of these pairs. Our discussions with subjects provided no indication that they interpreted the instruction to judge probability as an instruction to evaluate surprise. Furthermore, the surprise interpretation does not apply to the conjunction fallacy observed in judgments of frequency. We conclude that surprise and informational value do not properly explain the conjunction fallacy, although they may well contribute to the ease with which it is induced and to the difficulty of eliminating it.

Cognitive Illusions

Our studies of inductive reasoning have focused on systematic errors because they are diagnostic of the heuristics that generally govern judgment and inference. In the words of Helmholtz (1881/1903), "It is just those cases that are not in accordance with reality which are particularly instructive for discovering the laws of the processes by which normal perception originates." The focus on bias and illusion is a research strategy that exploits human error, although it neither assumes nor entails that people are perceptually or cognitively inept. Helmholtz's position implies that perception is not usefully analyzed into a normal process that produces accurate percepts and a distorting process that produces errors and illusions. In cognition, as in perception, the same mechanisms produce both valid and invalid judgments. Indeed, the evidence does not seem to support a "truth plus error" model, which assumes a coherent system of beliefs that is perturbed by various sources of distortion and error. Hence, we do not share Dennis Lindley's optimistic opinion that "inside every incoherent person there is a coherent one trying to get out," (Lindley, Note 3) and we suspect that incoherence is more than skin deep (Tversky & Kahneman, 1981).

It is instructive to compare a structure of beliefs about a domain, (e.g., the political future of Central America) to the perception of a scene (e.g., the view of Yosemite Valley from Glacier Point). We have argued that intuitive judgments of all relevant marginal, conjunctive, and conditional probabilities are not likely to be coherent, that is, to satisfy the constraints of probability theory. Similarly, estimates of distances and angles in the scene are unlikely to satisfy the laws of geometry. For example, there may be pairs of political events for which $P(A)$ is judged greater than $P(B)$ but $P(A/B)$ is judged less than $P(B/A)$—see Tversky and Kahneman (1980). Analogously, the scene may contain a triangle ABC for which the A angle appears greater than the B angle, although the BC distance appears to be smaller than the AC distance.

The violations of the qualitative laws of geometry and probability in judgments of distance and likelihood have significant implications for the interpretation and use of these judgments. Incoherence sharply restricts the inferences that can be drawn from subjective estimates. The judged ordering of the sides of a triangle cannot be inferred from the judged ordering of its angles, and the ordering of marginal probabilities cannot be deduced from the ordering of the respective conditionals. The results of the present study show that it is even unsafe to assume that $P(B)$ is bounded by $P(A\&B)$. Furthermore, a system of judgments that does not obey the conjunction rule cannot be expected to obey more complicated principles that presuppose this rule, such as Bayesian updating, external calibration, and the maximization of expected utility. The presence of bias and incoherence does not diminish the normative force of these principles, but it reduces their usefulness as descriptions of behavior and hinders their prescriptive applications. Indeed, the elicitation of unbiased judgments and the reconciliation of incoherent assessments pose serious problems that presently have no satisfactory solution (Lindley, Tversky & Brown, 1979; Shafer & Tversky, Note 4).

The issue of coherence has loomed larger in the study of preference and belief than in the study of perception. Judgments of distance and angle can readily be compared to objective reality and can be replaced by objective measurements when accuracy matters. In contrast, objective measurements of probability are often unavailable, and most significant choices under risk require an intuitive evaluation of probability. In the absence of an objective criterion of validity, the normative theory of judgment under uncertainty has treated the coherence of belief as the touchstone of human rationality. Coherence has also been assumed in many descriptive analyses in psychology, economics, and other social sciences. This assumption is attractive because the strong normative appeal of the laws of probability makes violations appear implausible. Our studies of the conjunction rule show that normatively inspired theories that assume coherence are descriptively inadequate, whereas psychological analyses that ignore the appeal of normative rules are, at best, incomplete. A comprehensive account of human judgment must reflect the tension between compelling logical rules and seductive nonextensional intuitions.

Reference Notes

1. Lakoff, G. *Categories and cognitive models* (Cognitive Science Report No. 2). Berkeley: University of California, 1982.
2. Beyth-Marom, R. *The subjective probability of conjunctions* (Decision Research Report No. 81-12). Eugene, Oregon: Decision Research, 1981.
3. Lindley, Dennis, Personal communication, 1980.
4. Shafer, G., & Tversky, A. *Weighing evidence: The design and comparisons of probability thought experiments.* Unpublished manuscript, Stanford University, 1983.

References

Ajzen, I. Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology,* 1977, *35,* 303–314.

Bar-Hillel, M. On the subjective probability of compound events. *Organizational Behavior and Human Performance,* 1973, *9,* 396–406.

Bruner, J. S. On the conservation of liquids. In J. S. Bruner, R. R. Olver, & P. M. Greenfield, et al. (Eds.), *Studies in cognitive growth.* New York: Wiley, 1966.

Chapman, L. J., & Chapman, J. P. Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology,* 1967, *73,* 193–204.

Cohen, J., & Hansel, C. M. The nature of decision in gambling: Equivalence of single and compound subjective probabilities. *Acta Psychologica,* 1957, *13,* 357–370.

Cohen, L. J. *The probable and the provable.* Oxford, England: Clarendon Press, 1977.

Dempster, A. P. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics,* 1967, *38,* 325–339.

Edwards, W. Comment. *Journal of the American Statistical Association,* 1975, *70,* 291–293.

Einhorn, H. J., & Hogarth, R. M. Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology,* 1981, *32,* 53–88.

Gati, I., & Tversky, A. Representations of qualitative and quantitative dimensions. *Journal of Experimental Psychology: Human Perception and Performance,* 1982, *8,* 325–340.

Goldsmith, R. W. Assessing probabilities of compound events in a judicial context. *Scandinavian Journal of Psychology,* 1978, *19,* 103–110.

Good, I. J. The probabilistic explication of information, evidence, surprise, causality, explanation, and utility. In V. P. Godambe & D. A. Sprott (Eds.), *Foundations of statistical inference: Proceedings on the foundations of statistical inference.* Toronto, Ontario, Canada: Holt, Rinehart & Winston, 1971.

Grice, H. P. Logic and conversation. In G. Harman & D. Davidson (Eds.), *The logic of grammar.* Encino, Calif.: Dickinson, 1975.

Hammond, K. R., & Brehmer, B. Quasi-rationality and distrust: Implications for international conflict. In L. Rappoport & D. A. Summers (Eds.), *Human judgment and social interaction.* New York: Holt, Rinehart & Winston, 1973.

Helmholtz, H. von. *Popular lectures on scientific subjects* (E. Atkinson, trans.). New York: Green, 1903. (Originally published, 1881.)

Jennings, D., Amabile, T., & Ross, L. Informal covariation assessment. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press, 1982.

Johnson-Laird, P. N., & Wason, P. C. A theoretical analysis of insight into a reasoning task. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking.* Cambridge, England: Cambridge University Press, 1977.

Jones, G. V. Stacks not fuzzy sets: An ordinal basis for prototype theory of concepts. *Cognition,* 1982, *12,* 281–290.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.) *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press, 1982.

Kahneman, D., & Tversky, A. Subjective probability: A judgment of representativeness. *Cognitive Psychology,* 1972, *3,* 430–454.

Kahneman, D., & Tversky, A. On the psychology of prediction. *Psychological Review,* 1973, *80,* 237–251.

Kahneman, D., & Tversky, A. On the study of statistical intuitions. *Cognition,* 1982, *11,* 123–141. (a)

Kahneman, D., & Tversky, A. Variants of uncertainty. *Cognition,* 1982, *11,* 143–157. (b)

Kirkwood, C. W., & Pollock, S. M. Multiple attribute scenarios, bounded probabilities, and threats of nuclear theft. *Futures,* 1982, *14,* 545–553.

Kyburg, H. E. Rational belief. *The Behavioral and Brain Sciences,* in press.

Lindley, D. V., Tversky, A., & Brown, R. V. On the reconciliation of probability assessments. *Journal of the Royal Statistical Society,* 1979, *142,* 146–180.

Mervis, C. B., & Rosch, E. Categorization of natural objects. *Annual Review of Psychology,* 1981, *32,* 89–115.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z., The use of statistical heuristics in everyday inductive reasoning. *Psychological Review,* 1983, *90,* 339–363.

Nisbett, R., & Ross, L. *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, N.J.: Prentice-Hall, 1980.

Osherson, D. N., & Smith, E. E. On the adequacy of prototype theory as a theory of concepts. *Cognition,* 1981, *9,* 35–38.

Osherson, D. N., & Smith, E. E. Gradedness and conceptual combination. *Cognition,* 1982, *12,* 299–318.

Rosch, E. Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization.* Hillsdale, N.J.: Erlbaum, 1978.

Rubinstein, A. False probabilistic arguments vs. faulty intuition. *Israel Law Review,* 1979, *14,* 247–254.

Saks, M. J., & Kidd, R. F. Human information processing and adjudication: Trials by heuristics. *Law & Society Review,* 1981, *15,* 123–160.

Shafer, G. *A mathematical theory of evidence.* Princeton, N.J.: Princeton University Press, 1976.

Slovic, P., Fischhoff, B., & Lichtenstein, S. Cognitive processes and societal risk taking. In J. S. Carroll & J. W. Payne (Eds.), *Cognition and social behavior.* Potomac, Md.: Erlbaum, 1976.

Smith, E. E., & Medin, D. L. *Categories and concepts.* Cambridge, Mass.: Harvard University Press, 1981.

Suppes, P. Approximate probability and expectation of gambles. *Erkenntnis,* 1975, *9,* 153-161.

Tversky, A. Features of similarity. *Psychological Review,* 1977, *84,* 327-352.

Tversky, A., & Kahneman, D. Belief in the law of small numbers. *Psychological Bulletin,* 1971, *76,* 105-110.

Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology,* 1973, *5,* 207-232.

Tversky, A., & Kahneman, D. Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in social psychology.* Hillsdale, N.J.: Erlbaum, 1980.

Tversky, A., & Kahneman, D. The framing of decisions and the psychology of choice. *Science,* 1981, *211,* 453-458.

Tversky, A., & Kahneman, D. Judgments of and by representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press, 1982.

Wyer, R. S., Jr. An investigation of the relations among probability estimates. *Organizational Behavior and Human Performance,* 1976, *15,* 1-18.

Zadeh, L. A. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems,* 1978, *1,* 3-28.

Zadeh, L. A. A note on prototype theory and fuzzy sets. *Cognition,* 1982, *12,* 291-297.

Zentner, R. D. Scenarios, past, present and future. *Long Range Planning,* 1982, *15,* 12-20.

Received July 7, 1982
Revision received May 2, 1983 ■

Third Edition of the *Publication Manual*

APA has just published the third edition of the *Publication Manual.* This new edition replaces the 1974 second edition of the *Manual.* The new *Manual* updates APA policies and procedures and incorporates changes in editorial style and practice since 1974. It amplifies and refines some parts of the second edition, reorganizes other parts, and presents new material. (See the March issue of the *American Psychologist* for more on the third edition.)

All manuscripts to be published in the 1984 volumes of APA's journals will be copy edited according to the third edition of the *Manual.* Therefore, manuscripts being prepared now should be prepared according to the third edition. Beginning in 1984, submitted manuscripts that depart significantly from third edition style will be returned to authors for correction.

The third edition of the *Publication Manual* is available for $12 for members of APA and $15 for nonmembers. Orders of $25 or less must be prepaid. A charge of $1.50 per order is required for shipping and handling. To order the third edition, write to the Order Department, APA, 1400 N. Uhle Street, Arlington, VA 22201.