ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES 37, 230–253 (1986)

Conjunction Errors: Evidence for Multiple Judgment Procedures, Including "Signed Summation"

J. FRANK YATES AND BRUCE W. CARLSON

The University of Michigan

A conjunction error is a judgment that a conjunctive event is more likely than one of the marginal events comprising the conjunction. Previous research has demonstrated conjunction errors in situations in which marginal events arose from a common generating process. In Studies 1–3, conjunction errors were reliably induced in cases where marginal events resulted from unrelated processes. This suggests that conjunction errors can be produced by "formalistic" judgment procedures, by which probability judgments for conjunctions are derived from probability judgments for marginal events according to some combination rule. Protocols from Studies 2 and 3 also indicated that subjects use multiple judgment procedures, several of which are capable of yielding conjunction errors. One new procedure suggested by the protocols is formalized in a "signed sum model." In this model, an event's likelihood is represented by a positive number, a negative number, or zero, depending on whether the event is considered likely, unlikely, or neither, respectively. The likelihood value for a conjunction is the sum of those for the constituent marginal events. In Study 4, this model successfully predicted when some subjects made zero, one, or two conjunction errors.

Imagine a person is presented with the following judgment problem: Consider arbitrary events A and B. Which is more likely to occur, event B or event [A & B], the conjunction of events A and B?

The conjunction rule of elementary probability theory requires that event [A & B] cannot be more likely than either event A or event B, i.e., $P(\text{A \& B}) \leq P(\text{A})$ and $P(\text{A \& B}) \leq P(\text{B})$. This conclusion is a consequence of what Tversky and Kahneman (1983) describe as the "extension rule" of probability. The extension rule says that any event is at least as probable as any other event it contains. The rule applies in the present situation because the conjunctive event [A & B] is contained both in event A and in event B.

Tversky and Kahneman (1982, 1983) have brought attention to instances in which humans' qualitative likelihood judgments reliably vio-

It is our pleasure to acknowledge the assistance of Halimah Hassan and Michael Fried in conducting the studies reported here. We have also benefited from criticisms of earlier drafts of this paper provided by members of the Michigan Decision Psychology Seminar. The comments of David Ronis and Shawn Curley were particularly helpful, as were those of two anonymous reviewers. Requests for reprints should be sent to J. Frank Yates, 136 Perry Building, Department of Psychology, The University of Michigan, Ann Arbor, MI 48104.

0749-5978/86 $3.00

late the conjunction rule. That is, their subjects commonly made statements of the form [A & B] $>_L$ B, where $>_L$ denotes a subjective qualitative likelihood relation, meaning that event [A & B] is considered more likely than event B. We will refer to a judgment that a conjunction is more likely than one of its constituent marginal events as a "conjunction error."[1]

One of Tversky and Kahneman's (1982) problems illustrates how conjunction errors often have been demonstrated. Subjects were given the following personality sketch of a hypothetical person (p. 92): "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." Subjects were then asked to rank order eight statements about Bill by their probability. Included in the list of statements were A = "Bill is an accountant," J = "Bill plays jazz for a hobby," and [A & J] = "Bill is an accountant who plays jazz for a hobby." (The subjects presumably did not see event labels such as "A" and "A & J.") Many of the subjects ranked the events A $>_L$ [A & J] $>_L$ J.

Why do people commit conjunction errors? This is the general issue considered in the research reported here. Previous results suggest several hypotheses.

One potential explanation of conjunction errors is that people are unaware of the conjunction rule. If this explanation has validity, then prior exposure to statistical training should decrease the incidence of conjunction errors. Evidence on the issue is mixed. In some tests, Tversky and Kahneman (1982, p. 93) found that statistical training had negligible effects on conjunction errors. In others (Tversky & Kahneman, 1983, p. 300), they discovered that subjects who had had several courses in statistics were much less likely to make conjunction errors.

A second hypothesis is that people who are aware of the conjunction rule do not accept its propriety. Tversky and Kahneman (1982) tested this hypothesis, too. Subjects who had responded to the above "Bill problem" were presented with the argument that "the probability that Bill is both an accountant and a jazz player cannot exceed the probability

[1] Tversky and Kahneman (1982) initially described any violation of the conjunction rule as a "conjunction effect." More recently (Tversky & Kahneman, 1983), they have applied the term "conjunction fallacy" to the occurrence of such a violation in a within-subjects design. They have used the expression "conjunction error" to describe conjunction rule violations revealed in between-subjects designs. We have found it easier to follow the convention taken by Morier and Borgida (1984). These authors appear to apply the term "conjunction error" to any specific violation of the conjunction rule. They reserve "conjunction fallacy" for the general phenomenon whereby such errors take place.

that he is a jazz player, because every member of the former category is also a member of the latter" (p. 95). Subjects who committed conjunction errors generally agreed that their judgments were mistakes, because they conflicted with the conjunction rule. Another test was not so favorable to the acceptance hypothesis. Tversky and Kahneman (1983, pp. 299–300) asked subjects to compare the convincingness of two arguments concerning the probability ordering of a conjunction and a constituent marginal event. One argument was essentially the same as the above extensional one. The other defended conjunction errors on the basis of the representativeness heuristic (Kahneman & Tversky, 1972). A majority of the subjects found the representativeness argument more compelling.

Markus and Zajonc (in press) propose that conjunction errors result from the subject misunderstanding what the events are whose chances of occurrence are being requested. For example, in the Bill problem, the presence of the event [A & J] = "Bill is an accountant who plays jazz for a hobby" may lead to the misinterpretation of the event J = "Bill plays jazz for a hobby" as the event [A′ & J] = "Bill is a nonaccountant who plays jazz for a hobby," where the prime on A′ indicates the complement of event A. Although [A & J] $>_L$ J necessarily violates probability theory, the judgment [A & J] $>_L$ [A′ & J] does not.

Tversky and Kahneman (1983, p. 299) asked subjects to rate the probability of events of the form [A & B] and B. They found that the ratings violated the conjunction rule, too, i.e., rating (A & B) > rating (B). This result was taken as evidence against the misunderstanding hypothesis. The apparent reasoning behind this conclusion is that, since the events were rated independently of each other, the event B should not have been misinterpreted as the event [A′ & B].

Morier and Borgida (1984) tested the misunderstanding hypothesis more directly. This was done by asking subjects to rank order the probabilities of events in a list that included not only those of the form A, B, and [A & B], but an explicit statement of event [A′ & B], too. Seeing event [A′ & B] in the list, the subject should not then confuse it with event B. Morier and Borgida found that the inclusion of [A′ & B] reduced the incidence of conjunction errors in one of two judgment problems presented to subjects, but not the other. So, again, there is mixed support for the proposed explanation for conjunction errors.

Judgment research over the past several years has shown that people make likelihood judgments according to a wide variety of procedures (cf. Einhorn & Hogarth, 1981; Pitz & Sachs, 1984; Rapaport & Wallsten, 1972; Slovic, Fischhoff, & Lichtenstein, 1977). Some of those procedures allow conjunction errors; others do not. The explanation for conjunction errors proposed by Tversky and Kahneman (1983) can be viewed in these terms. Specifically, these authors hypothesize that people sometimes make judgments by procedures that inadvertently permit conjunction errors to happen.

Tversky and Kahneman (1983, p. 294) indicate that "*natural assessments* are routinely carried out as part of the perception of events. . . . Such natural assessments include computations of similarity and representativeness, attributions of causality, and evaluations of the availability of associations and exemplars." Further, they argue (p. 310) that "a judgment of probability or frequency is commonly biased toward the natural assessment that the problem evokes. . . . These assessments are not constrained by the extension rule." In particular, judgments made according to the representativeness and availability (Tversky & Kahneman, 1973) heuristics are natural assessments that can yield conjunction errors. For example, in the Bill problem, "Bill is an accountant who plays jazz for a hobby" is judged more likely than "Bill plays jazz for a hobby" because the former event is more representative of Bill's personality sketch than the latter. Tversky and Kahneman (1983) offer similar demonstrations of how judgments according to availability can produce conjunction errors.

There is another class of judgment procedures that, in principle, could account for some conjunction errors. What might be called "formalistic" procedures share the following characterization: The person's task requires that likelihood judgments be made for several events. Judgments for some of those events are made directly. That is, direct judgments do not rely upon any other likelihood judgments. However, judgments for the remaining events are derived—consciously or otherwise—from judgments for other events according to combination rules accepted by the person. Judgment via probability theory would be an example of a formalistic procedure. Suppose, for instance, that a probability judgment for the event [A & B] is required. Assuming that events A and B are independent, the desired judgment can be derived from judgments for the events A and B: $P'(A \& B) = P'(A)P'(B)$, where the notation P' indicates a probability judgment.[2]

Shanteau (1975) and Troutman and Shanteau (1977) have demonstrated that some probability judgments can be well described by the formalistic procedure of averaging. In a study more directly relevant to the present problem, Wyer (1976) asked subjects to report probability judgments for marginal, conditional, and conjunctive events, e.g., A, B, A|B, and [A & B]. He indicated (p. 9) that, "Indeed, estimates of P_{AB} were less

[2] The term "probability judgment" rather than "subjective probability" is used to emphasize that the person's judgments may or may not satisfy all the probability axioms. "Subjective probability" implies that the axioms *are* satisfied, the very issue being contested here.

than the estimates of both component probabilities (e.g., both P_A and P_B) in only a minority of cases.'' Thus, Wyer obtained clear evidence of conjunction errors in quantitative likelihood judgment. Wyer also found that the best models for his subjects' responses included both a multiplying component (consistent with probability theory) and an averaging component. The averaging component requires that, if event A is judged more likely than event B, then $P'(A) > P'(A \& B) > P'(B)$. Virtually all reported instances of conjunction errors, e.g., in the Bill problem, have involved situations in which one of the marginal events seems highly probable while the other appears remote. Consistent with averaging, subjects generally have judged the conjunction to be more likely than one of the constituent marginal events, but not the other.

This brief review has indicated that conjunction errors are compatible with several previously proposed explanations. It is conceivable that there is a single mechanism that fully accounts for all conjunction errors. This appears doubtful, however. It is more plausible that, on different occasions, conjunction errors occur for each of the reasons that have been offered. There may well be other sources of conjunction errors, too. So a legitimate goal is to work toward completing the catalog of *documented* conjunction error explanations. Such was the aim of the studies reported here.

STUDY 1

The evidence for conjunction errors arising from natural assessments is convincing. However, are there conjunction errors that can be explained by formalistic procedures, but not natural assessments? Although Wyer's (1976) data are consistent with an averaging explanation of conjunction errors, they are also compatible with an account according to natural assessments. An approach for comparing these explanations is implicit in an observation made by Tversky and Kahneman (1983). These authors note that judgments of the representativeness (and hence probability) of a conjunction of events A and B are affected by the compatibility of the two events. As a corollary, they submit 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. (p. 305)

Suppose that the constituent marginal events in a conjunction are generated by processes that are unrelated to each other. Then it seems implausible that natural assessments could lead to conjunction errors. On the other hand, formalistic procedures, including operations equivalent to averaging, could. It so happens that the marginal events in the conjunctions studied by Wyer (1976) involved common rather than unrelated processes. In one type of conjunction, for example, one marginal event was that a person possessed a certain gene, while the other was that that individual exhibited a particular characteristic. The present study tested for the occurrence of conjunction errors in situations in which marginal events did indeed appear to rest upon unrelated processes.

Method

Subjects

The 78 subjects in the study were students in introductory psychology courses at the University of Michigan. They participated in the study as a course credit option; they were not paid.

Materials

Five judgment problems were presented in questionnaire form. Descriptions of the marginal events in each problem, denoted by A and B, are contained in Table 1.

In each problem, the task was to rank order several events from most to least likely, using 1 for the most probable event, 2 for the second-most probable event, and so on. The McEnroe problem was introduced with the statement, "Suppose John McEnroe plays in the first round of the

TABLE 1
FOCAL MARGINAL EVENTS IN LIKELIHOOD JUDGMENT PROBLEMS: STUDY 1

Problem	Event A	Event B
McEnroe	McEnroe will win the match	McEnroe will lose the first set
Reagan	Reagan will cut federal support to local government	Reagan will provide federal support for unwed mothers
Blanchard/ Derek	Governor Blanchard will succeed in raising the Michigan state income tax	Bo Derek will win an Academy Award for the movie that she is currently making
Regents/ Floyd	The University of Michigan regents will increase tuition next year	Ray Floyd will win the U.S. Open golf tournament
Urn	The marble you draw from the second jar is black	The marble you draw from the first jar is red

Wimbledon tennis tournament in June of this year." Afterward, the subject was asked to rank order in likelihood the events whose descriptions followed. The Reagan, Blanchard/Derek, and Regents/Floyd problems were all introduced with the simple request, "Please rank order the following events by their probability of occurrence in 1983." The introduction to the Urn problem was the following:

> Suppose that two jars containing different colored marbles are set on a table before you. The first jar contains 5 red and 20 blue marbles. The second jar contains 80 black and 10 white marbles. Without looking, you draw one marble from each jar.

The subject was then instructed to rank in order of likelihood the events described thereafter. Included in the list of events to be rank ordered in each problem were A, B, and their conjunction, [A & B]. The event lists for the McEnroe, Reagan, Blanchard/Derek, and Regents/Floyd problems also contained a fourth event.

The McEnroe and Reagan problems were adapted directly from problems introduced by Tversky and Kahneman (1982). In the original version of the McEnroe problem, the tennis player was Bjorn Borg, who has since retired. Tversky and Kahneman (1982, p. 96) described the conjunction in the tennis problem as "Borg will lose the first set but win the match," using the connective *but* rather than the more conventional *and,* which *was* used in their version of the Reagan problem. To determine whether that distinction makes a difference, two forms of the McEnroe and Reagan problems were used in the present study, one containing the *but* connective, the other the *and* connective. The *and* connective was employed in all other problems.

In contrast to the McEnroe and Reagan problems, the Blanchard/Derek and Regents/Floyd problems concerned marginal events we judged to be generated from unrelated rather than common processes. Michigan fiscal politics appeared to us as remote from film making. Similarly, it was difficult for us to see how a common process could account for the setting of University of Michigan tuition rates and the outcome of the U.S. Open golf tournament.

All the problems were designed such that—as in the McEnroe and Reagan problems—most subjects would be expected to judge $A >_L B$. At the time the study was conducted, William Blanchard, Governor of the State of Michigan, was lobbying hard for an increase in the state income tax to relieve a budget crisis. His success was heavily predicted by newspaper and television political analysts. Most students were thought to be aware that the University of Michigan's regents had increased tuition practically every year in the previous decade, and most observers felt they would continue to do so. The chances of Bo Derek winning an Academy Award and of Ray Floyd winning the U.S. Open golf tournament were expected to be seen as slight in part because of the tremendous competition for those prizes.

The selection of balls from an urn, as in the Urn problem, is a transparently random process. The Urn problem also admits a readily accessible extensional cognitive representation that should facilitate a proper ordering of the events, assuming that people understand the principle implicit in the conjunction rule. So responses to the Urn problem should have provided a point of reference for responses to the other problems.

Procedure

Subjects were assigned to small group sessions (1–15 subjects) at the convenience of their schedules. The McEnroe, Reagan, Blanchard/Derek, and Regents/Floyd problems were presented in a "Part 1" questionnaire. The McEnroe and Reagan problems were presented first as a randomized pair. An intermediate task not relevant to the present study followed the Blanchard/Derek and Regents/Floyd problems. After all the Part 1 materials were collected, the Urn problem was presented in a "Part 2" questionnaire. At all stages of the procedure, subjects were allowed and encouraged to ask questions about the tasks being requested of them. A postexperimental questionnaire asked the subject if he or she had ever taken a course in which probability had been taught.

Results

No significant presentation order effects were observed. The subjects' responses were also unaffected by their prior exposure to probability concepts. Another negative finding was that whether the connective in the McEnroe and Reagan problems was *and* or *but* did not influence the incidence of conjunction errors. Accordingly, the summary of results described below ignores that distinction.

Table 2 presents the percentages of judged likelihood ordering patterns found in Study 1. These percentages are easily demonstrated to result from nonrandom responding.

The Total conjunction error row of Table 2 shows, for each problem, the percentage of all patterns which violate the conjunction rule. According to Cochran's (1950) test for comparisons of percentages in matched samples, the incidence of conjunction rule violations was significantly lower for the Urn problem than for the other problems ($Q(1) = 31.78$, $p < .001$). A significant majority of the subjects made no conjunction errors on the Urn problem ($p < .001$). In contrast, for each of the other problems a majority of the subjects made at least one conjunction error, although the only majority that was statistically significant was in the McEnroe problem ($p < .001$). Of particular importance is the fact that the total incidences of conjunction errors for the non-Urn problems dif-

TABLE 2
PERCENTAGES OF JUDGED LIKELIHOOD ORDERING PATTERNS: STUDY 1

Likelihood ordering patterns	Problem: Common process — McEnroe	Common process — Reagan	Unrelated processes — Blanchard/Derek	Unrelated processes — Regents/Floyd	Random processes — Urn
Single conjunction error:					
$A \geqslant_L A\&B >_L B$	47.4	42.3	55.1[a]	53.8	28.2[b]
$B >_L A\&B >_L A$	12.8	3.9	0.0	0.0	0.0
Double conjunction error:					
$A\&B >_L A, B$	10.2	11.5	1.3	1.3	1.3
Total conjunction error	70.5	57.7	56.4	55.1	29.5
No conjunction error	29.5	42.3	43.6	44.9	70.5

Note. $N = 78$.
[a] Includes one case in which $A =_L A\&B >_L B$, where $=_L$ denotes judged equal likelihood.
[b] Includes three cases in which $A =_L A\&B >_L B$.

fered from one another only marginally significantly ($Q(3) = 6.73$, $.05 < p < .10$), even though double conjunction errors (i.e., $[A \& B] >_L A$ and $[A \& B] >_L B$) were more prevalent for the McEnroe and Reagan problems than all others ($Q(1) = 17.76$, $p < .01$). The frequency of conjunction errors in Study 1 was much less than that reported by Tversky and Kahneman (1982). It is not obvious why this was so.

The Single conjunction error rows of Table 2 present the percentages of event orderings involving a single conjunction error. These include the patterns one would expect if subjects made judgments according to an averaging type of procedure. If, as anticipated, subjects considered event A more likely than event B, then averaging would predict the ordering $A >_L [A \& B] >_L B$. This pattern was much more common for the non-Urn problems than for the Urn problem ($Q(1) = 15.06$, $p < .001$). The non-Urn problems did not differ among themselves in the incidence of that pattern ($Q(3) = 4.34$, ns).

Discussion

The comparative paucity of conjunction errors on the Urn problem suggests that many people do indeed have an understanding of the conjunction rule. Apparently, however, they are much less likely to apply it to problems with more naturalistic content; instead, they rely on judgment procedures that sometimes conflict with the conjunction rule. As Tversky and Kahneman (1983) noted, this is unusual, since natural rather than abstract problem content often enhances performance on reasoning tasks (e.g., Gilhooly & Falconer, 1974).

Study 1 demonstrated that conjunction errors are almost as common in judgment problems in which marginal events rest on unrelated processes as they are in problems involving common processes. Considering also the fact that a substantial minority of subjects did commit conjunction errors on the Urn problem, the results of Study 1 indicate that procedures involving natural assessments are not the only ones that cause people to make conjunction errors. Formalistic procedures, exemplified by the averaging of quantitative probability judgments, seem to be plausible candidates as the source of some of those errors.

STUDIES 2 AND 3

Studies 2 and 3 were designed to gather additional data on the incidence of conjunction errors involving marginal events resting on both common and unrelated processes. Study 1 provided indirect evidence for formalistic judgment procedures leading to conjunction errors. And although that evidence is consistent with averaging as the particular formalistic procedure yielding conjunction errors, the possibility exists that there are other such procedures. So Studies 2 and 3 also served an exploratory function. The plan was to have subjects articulate the reasoning they use when they solve judgment problems similar to those typically presented in studies of conjunction errors. The resulting protocols could then be examined for clues as to the judgment procedures that cause people to commit such errors. These suggested procedures could then be subjected to more rigorous verification testing by experimental means.

Method

Subjects

Forty-six students in an upper level undergraduate psychology course at the University of Michigan participated in Study 2 as an exercise used in introducing the conjunction error phenomenon. They were not paid. Thirty-three University of Michigan students served as paid subjects in Study 3. They were recruited from a departmental subject pool.

Materials

Each subject considered six judgment problems presented in a questionnaire. As in Study 1, each problem required the subject to rank order four events by their probability of occurrence, using the number 1 for the most likely event through 4 for the least likely event. The four events included two marginal events, which can be denoted by A and B. Event A was expected to be considered highly likely by most subjects; event B was expected to be judged quite unlikely. In half the problems events A and B were generated by what we assumed to be common processes; they arose from unrelated processes in the other problems. The third and fourth events in a given problem were the conjunctive events [A & B] and [A′ & B′]. For a given problem, the marginal events were described first, then the conjunctive events. For half the problems event A appeared before event B; conversely for the other problems. There was a similar counterbalancing in the presentation orders for the two conjunctive events. The Appendix contains descriptions of events A and B in the judgment problems considered by the subjects.

Procedure

Subjects in Study 2 participated in a single group session, although they were required to complete their tasks independently. Instructions were presented orally as well as in each subject's questionnaire. Subjects were encouraged to ask questions if they felt they did not understand their task. Besides rank ordering the probabilities of the events in each problem, the subject was instructed: "After you have ranked the events, you should *briefly,* i.e., *in no more than 5–8 lines,* indicate why you ranked the events the way you did."

Subjects in Study 3 participated in individual, tape-recorded sessions. After receiving general instructions about the task of rank ordering events in terms of their probability, the subject was also asked to read each problem out loud. This was intended to ensure that the subject understood each problem correctly. The subject was also told the following:

> As you are thinking about your solution to the problem, say what is going through your mind. In other words, *think out loud.* What's your reasoning, your rationale? Don't worry about whether or not something sounds right or sounds silly. You should speak your thoughts *as they occur to you.* Don't hold back until *after* you have made up your mind.

This concurrent reporting procedure was intended to counteract potential memory-related biases (Ericsson & Simon, 1980; Nisbett & Wilson, 1977). The final instruction was for the subject to write down his or her judgments as they were decided.

Results and Discussion

Conjunction Error Frequencies

Table 3 presents the percentages of conjunction errors observed in the various judgment problems. The relevant rows of the table are those indi-

TABLE 3
JUDGMENT PROBLEM CONJUNCTION ERROR PERCENTAGES, STUDIES 2–4

Ensemble	Processes	Study	Chances	Percentage conjunction errors[c,d]		
				0	1	2
McEnroe	C[a]	4	LL[b]	43.1****	21.6	35.3
		2/3	LU	36.7	59.5	3.8
		4	UU	62.0	28.0	10.0
Football/Mideast	U	4	LL	30.0****	40.0	30.0
		2/3	LU	31.6	68.4	0.0
		4	UU	51.9	44.2	3.8
Reagan	C	2/3	LU	31.6****	68.4	0.0
		4	UU	53.9	26.5	19.6
"60 Minutes"/ mortgages	U	4	LL	51.5****	21.8	26.7
		3	LU	42.4	54.5	3.0
USSR	C	4	LL	34.3****	8.8	56.9
		2/3	LU	24.7	75.3	0.0
Weather/AIDS[e]	U	2	LU	48.9***	51.1	0.0
		3	LU	27.3****	57.6	15.2
		4	UU	64.7	25.5	9.8
Congress/Derek	U	2	LU	53.3	42.2	4.4
Election/ "Breakin' "	U	4	LL	34.0***	36.0	30.0
		4	UU	58.8	37.3	3.9
Cars/Tigers	U	4	LL	47.1*	19.6	33.3
		4	UU	48.0	36.0	16.0

[a] C—common processes; *U*—unrelated processes.
[b] LL—event A and event B likely; LU—event A likely, event B unlikely; UU—event A and event B unlikely.
[c] Percentages sometimes do not sum to 100 due to rounding error.
[d] Significance tests and levels apply to across-problem percentages within ensembles.
[e] Significance tests and levels involve Study 2 versus Study 4 and Study 3 versus Study 4.
* $p < .10$.
** $p < .05$.
*** $p < .01$.
**** $p < .001$.

cated in the Study column to refer to either Study 2, Study 3, or both. The designation "LU" in the Chances column is a reminder that, in each case, the subject was expected to consider event A likely and event B unlikely. As in Study 1, conjunction errors occurred just about equally often when event-generating processes were unrelated as when they were common. Four of the problems considered by subjects in Studies 2 and 3 were identical in the two studies. Two involved common (C) processes, the others unrelated (U) processes. The difference in the incidence of conjunction errors across problems was not statistically significant ($Q(3) = 4.29$, ns).

Judgment Procedures

The subjects' written and tape-recorded protocols were examined for indications of the procedures they used in arriving at judgments about the likelihood of the conjunctive event [A & B] relative to the likelihoods of the marginal events A and B. The subjects' statements did not have sufficient clarity to always reliably distinguish when the subject was using one approach to the task rather than another. Moreover, the interpretation of protocols is always inescapably a subjective exercise. Nevertheless, there were still many instances in which subjects *did* appear to definitely employ one particular strategy. Responses in those instances allowed us to assemble a taxonomy of several procedures that at least some of the subjects used some of the time.

Following are the procedures and representative protocol excerpts illustrating them:

Extensional reasoning–concept. Some subjects' opinions agreed with the conjunction rule because their judgment procedures involved extensional reasoning. In some cases, this strategy appeared to be spontaneous, resting on a fundamental, untutored understanding of extensional principles (example: "3 [A & B], 4 [A′ & B′] have two requirements making them more difficult to fulfill" [McEnroe—Study 2—Subject 42]).

Extensional reasoning–probability theory. Other subjects considered conjunctions less likely than constituent marginal events because they applied probability theory rules recalled from their studies (example[3]: "I'm gonna give something . . . 90% probability to {B}. To {A} I'm gonna give 80% probability because in Europe . . . there are many missiles aimed to Russia. . . . While the combination of both events is gonna be 82 . . . no, 72%, the multiplication of {A} and {B}, that's why {A & B} is gonna be smaller" [USSR—Study 3—Subject 5]).

Marginal event misunderstanding. As hypothesized by Markus and Zajonc (in press), some subjects committed conjunction errors because they interpreted the marginal event B as the conjunction [A′ & B] (example: "{A & B} is higher than {B} because the only way he can get away w/ increasing military [B] is to increase something else [A]" [Reagan—Study 2—Subject 40]).

Representativeness. Some subjects' judgments appeared to rest on the

[3] Expressions in braces refer to the subject's literal statement interpreted in terms of the events labeled A and B.

representativeness heuristic (example: "The choices seem consistent with Reaganomics." [Note: Gave ranking A $>_L$ [A & B] $>_L$ [A′ & B′] $>_L$ B.] [Reagan—Study 2—Subject 39]).

Substantive reasoning. It was especially common for subjects to justify their ordering of marginal and conjunctive events on the basis of arguments involving the substance of the given situation. Sometimes those arguments led to judgments consistent with the conjunction rule, other times not (example: "John tends to win most matches, but losing in the first set could make him play harder, and intimidate his opponent even more" [McEnroe—Study 2—Subject 23]).

Averaging. Several subjects made judgments that appeared to rest on the reasoning implicit in averaging (example: "I am certain that UM will win at least four games, thus {A} = No. 1. I am so sure of this that even though I feel it is least likely that Syria and Israel will sign ({B} = No. 4) it carries enough weight so that the combined aspects are rated second-most likely ({A & B} = No. 2)" [Football/Mideast—Study 2—Subject 12]).

Signed summation. A new formalistic approach we have called "signed summation" seemed to emerge. This procedure occurs in two stages. In the first, the person classifies each marginal event as "likely," "unlikely," or neither. In the second stage, the person derives a judgment of the chances for the conjunction of the marginal events from the perceived chances for the marginal events. When conjoined, two likely (unlikely) events are seen as even more likely (unlikely) than either (example: "These two separate events (Congress [raising] taxes, & no Bo nomination) are almost guaranteed locks I believe. Putting them both in the same statement [A′ & B′] then *must* be rated w/ the highest prob" [Congress/Derek—Study 2—Subject 8]). The likelihood of a conjunction of a likely and an unlikely event is something of a compromise.

It is noteworthy that individual subjects often used several procedures, sometimes even in the same problem. One might expect that, if the subject happened to use extensional reasoning on one problem, then he or she would continue to use that strategy, evidencing a type of insight. This did not always happen. Some amount of insight might have taken place, however. Suppose a subject made no conjunction errors on Problem N in the series of six considered. Then the subject was significantly more likely to make no conjunction errors on Problem $N + 1$ than had he or she made at least one error on Problem N (45.0% vs 15.3%). This comparison assumes that at least one error had been made on Problem $N - 1$, i.e., that the insight had not occurred on the previous problem.

STUDY 4

The most surprising outcome of Studies 2 and 3 was the suggestion of a new formalistic approach to making likelihood judgments, the signed summation procedure. The procedure can be stated more formally as the following *signed sum model of qualitative likelihood judgment.*

The *qualitative likelihood* function λ is defined over the relevant family of events. The range of λ includes the entire real line. In particular, the *qualitative likelihood* of arbitrary event E, $\lambda(E)$, can be negative.

For events E and F, $\lambda(E) > \lambda(F)$ if and only if $E >_L F$, i.e., E is judged more likely than F. Similarly, $\lambda(E) = \lambda(F)$ if and only if $E =_L F$, i.e., E and F are judged equally likely. As a natural extension, $\lambda(E) \geq \lambda(F)$ is equivalent to the judgment that E is at least as likely as F, $E \geq_L F$.

An event E is defined as *likely* if it is judged more likely than its complement, i.e., $E >_L E'$. Likely events have positive qualitative likelihoods. That is, if E is likely, then $\lambda(E) > 0$.

Event E is said to be *unlikely* if it is considered less likely than its complement, i.e., $E' >_L E$. Unlikely events are assigned negative qualitative likelihoods. So if E is unlikely, then $\lambda(E) < 0$.

An *indifferent* event is defined as one that is neither likely nor unlikely. In other words, if E is indifferent, then it is judged to be just as likely to occur as to not occur: $E =_L E'$. The corresponding qualitative likelihoods would be zero, i.e., $\lambda(E) = \lambda(E') = 0$.

The *signed sum rule* specifies the relationship between the qualitative likelihood of a conjunction and the qualitative likelihoods of the constituent marginal events: For events E and F, with conjunction [E & F],

$$\lambda(E \,\&\, F) = \lambda(E) + \lambda(F). \tag{1}$$

This completes the statement of the signed sum model.

The signed sum model implies that conjunction errors *can* occur, but only under specified conditions. Figure 1 illustrates the model's predictions regarding the relative chances of conjunctions and their constituent marginal events. Suppose that events A and B are both considered likely to occur. The signed sum model says that their conjunction should be seen as especially likely to happen. In particular, as Case 1 in Fig. 1 suggests, the subject should make double conjunction errors, i.e., $[A \,\&\, B] >_L A$ and $[A \,\&\, B] >_L B$.

Now suppose that event A is judged likely, whereas event B is considered unlikely. Then, as illustrated in Cases 2–3 in Fig. 1, according to the signed sum model, the conjunction could be either likely or unlikely. The resolution of the conflict depends on the relative degrees of likeliness and unlikeliness of events A and B. If A is more likely than B is unlikely, then [A & B] will be likely; conversely, if the imbalance is in the opposite direction. In either case, the subject should make a single conjunction error, $[A \,\&\, B] >_L B$.

Finally, consider the circumstance in which both event A and event B

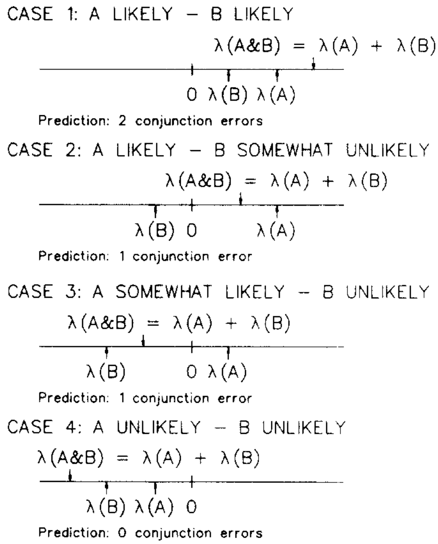

FIG. 1. Qualitative likelihoods and conjunction error predictions according to the signed sum model.

are seen as unlikely. The signed sum model prescribes that their conjunction should be thought very unlikely. As indicated in Case 4 of Fig. 1, the subject should not commit a conjunction error in such a situation.

Studies 1–3 support the conclusion that no single procedure is likely to account for all subjects' judgments of the relative chances of conjunctive and marginal events. Indeed, they indicate that even a single individual is unlikely to use only one judgment procedure all the time. However, the exploratory data of Studies 2 and 3 suggest that a procedure equivalent to the signed sum model is among the procedures at least some subjects use on some occasions. Study 4 was intended to test this model according to the qualitative predictions outlined above.

Method

Subjects

One hundred five subjects participated in Study 4. They were recruited from the general student and employee community of the University of Michigan and a local law office. The subjects were paid.

Materials

The basic judgment task of subjects in Study 4 was the same as that in Studies 2 and 3. That is, the subjects were required to rank order in judged probability four events of the form A, B, [A & B], and [A′ & B′].

The Appendix presents the particulars of the eight ensembles of problems considered by the subjects. The first six ensembles contained problems with the same subject matter as those in Studies 2 and 3. The last two ensembles, containing the Election/"Breakin' " and the Cars/Tigers problems, concerned topics not considered previously.

Recall that all the problems in Studies 2 and 3 were such that event A was expected to be seen as likely and event B as unlikely. So the responses of subjects in those studies could be used as part of the data for testing the signed sum model. Some of the additional problems added to each ensemble in Study 4 were such that both event A and event B should have been perceived as likely by most subjects. These problems can be distinguished from others by the notation "LL." Study 4 also included problems in which events A and B were both expected to be considered unlikely. Such problems can be characterized by "UU."

The Election/"Breakin' " and Cars/Tigers problems differed slightly from the others in the following way. Besides events A, B, and [A & B], the fourth event in these problems was the disjunction [A or B] rather than the conjunction [A′ & B′]. This manipulation was intended to test the robustness of the effects predicted by the signed sum model.

Two forms of the questionnaire containing the Study 4 problems were constructed. Each of the alternate forms contained either the LL or the UU form of each problem in each ensemble. The assignment of problems to forms was random, with the constraint that each form contained an equal number of LL and UU problems.

Procedure

Questionnaire forms were assigned to subjects at random. Subjects were given the opportunity to ask questions before completing the questionnaires. They were required to work alone.

Results and Discussion

Figure 2 contains a graphic display of the percentages of subjects who made zero, one, or two conjunction errors on the LL, LU, and UU problems in the McEnroe and in the Football/Mideast ensembles. These were the ensembles used to test all three of the signed sum model predictions described above. Zero conjunction errors were most common in UU problems as compared to LU and LL problems. Also, as expected, single conjunction errors were much more frequent in LU rather than LL and UU problems. Finally, double conjunction errors were numerous only in LL problems.

Table 3 contains the actual percentages of subjects making zero, one,

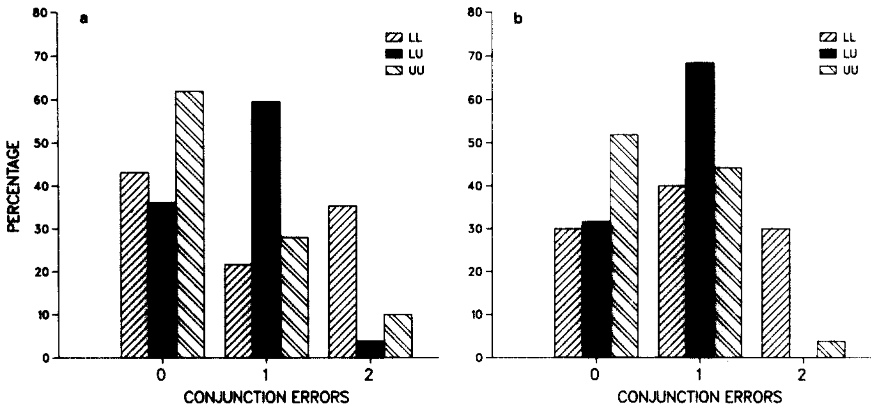

FIG. 2. Percentage of conjunction error frequencies by expected marginal event likelihood classification: (a) McEnroe ensemble; (b) Football/Mideast ensemble. LL—event A and event B likely; LU—event A likely, event B unlikely; UU—event A and event B unlikely.

and two conjunction errors in each of the problem ensembles. In all except the Weather/AIDS ensemble, negative significance tests indicated that subjects in Study 3 responded to particular LU problems the same way those in Study 2 did. That is why the error percentages for that problem are listed separately by study rather than pooled. For every ensemble, the pattern of error frequencies was as predicted by the signed sum model ($p < .01$ in all instances except one, by χ^2).

The signed sum model predicts that, when both event A and event B are likely, their conjunction [A & B] should be more likely than either of them. The present data suggest that, in many instances, [A & B] will be judged more likely than even the *disjunction* [A or B]. In the Election/"Breakin' " LL problem, 46% of the subjects considered the conjunction more likely than the disjunction. In the Cars/Tigers LL problem, the corresponding proportion was even higher, 54.9%. So the predicted effects appear to be quite robust.

GENERAL DISCUSSION

Several conclusions follow from the present studies. People make judgments about the relative chances of marginal and conjunctive events according to numerous different procedures. Some of those procedures admit conjunction errors, i.e., judgments that a conjunction is more likely than one of the constituent marginal events; others do not. There is no reason to doubt that, as previously demonstrated by Tversky and Kahneman (1982, 1983), some conjunction errors result from judgment according to "natural assessments," such as representativeness and availability. However, conjunction errors also arise from various other procedures, including "formalistic" ones.

One formalistic procedure some people use is "signed summation." This judgment method proceeds in two stages. First, the person determines the qualitative likelihoods of the two relevant marginal events, including whether each is "likely" or "unlikely." The qualitative likelihood of their conjunction is then derived from their individual qualitative likelihoods. This is done in a fashion equivalent to the summation of signed real numbers. Under specified conditions, this procedure yields zero, one, or two conjunction errors in a given judgment problem.

The signed sum model differs most notably from previously recognized judgment procedures in its categorization of events as either likely or unlikely, rather than simply locating them on a unidirectional continuum of certainty. The form of the signed sum model is different from an averaging model that involves only the events explicitly presented to the person. If events A and B are not considered equally likely, such a model would always predict a single conjunction error. The signed sum model and averaging agree, however, if certain strong assumptions are made.

Suppose the person has an initial opinion I about the chances that the conjunction of events A and B is going to happen, e.g., that it is twice as likely to occur as not occur. Let $s(I)$ denote a scale value for opinion I. Similarly, let $s(\mathrm{A})$ and $s(\mathrm{B})$ indicate the scale values for the person's opinions that events A and B will occur, with high values implying greater chances. The parameters $w(I)$, $w(\mathrm{A})$, and $w(\mathrm{B})$ can be taken as weights corresponding to the impact of the respective opinions on the person's judgment of how likely the conjunction [A & B] is. A rule similar to Anderson's (1965) weighted averaging model with an initial impression would characterize the scale value of [A & B] as follows:

$$s(\mathrm{A}\ \&\ \mathrm{B}) = \frac{w(I)s(I) + w(\mathrm{A})s(\mathrm{A}) + w(\mathrm{B})s(B)}{w(I) + w(\mathrm{A}) + w(\mathrm{B})}. \qquad (2)$$

Unlike an "ordinary averaging rule" involving only $s(\mathrm{A})$ and $s(\mathrm{B})$, this "initial impression averaging rule" does not require that $s(\mathrm{A}\ \&\ \mathrm{B})$ be between $s(\mathrm{A})$ and $s(\mathrm{B})$. In particular, it leads to the same type of event orderings as the signed sum rule, if saying a marginal event is "likely" means its chances of occurrence are considered better than the initial opinion of the conjunction's chances, and "unlikely" means the opposite.

So one critical distinction between signed summation and initial impression averaging is whether, when considering the likelihood of a marginal event, the person makes reference to an initial opinion about [A & B] rather than some absolute standard of what is "likely." The protocol data from Studies 2 and 3 suggest that the latter is more plausible. Future experiments intended to more rigorously differentiate the models will exploit this distinction. At the present time, the signed sum model is concerned only with qualitative, i.e., ordinal, likelihood judgments. Later the model will be extended in an attempt to account for quantitative opinions, e.g., probability judgments. It should be straightforward to experimentally contrast the initial impression averaging model from a quantitative version of the signed sum rule. This can be done using procedures involving three or more marginal events (cf. Anderson, 1965, 1981). That the currently available data do not sharply distinguish the signed sum model and the initial impression averaging model does not detract from the fact that either accounts for judgments that other models have difficulty explaining.

Within a given problem, the majority of conjunction error frequencies did not always conform to the predictions of the signed sum model (e.g., in the LL problems depicted in Fig. 2). However, we chose as our bases of comparison the numbers of conjunction errors subjects would make on different problem types, e.g., LL versus UU and LU. This seemed more meaningful than a comparison against a uniform distribution, since subjects do not normally make their judgments randomly. Nevertheless, it is clear that, although the model can explain some conjunction errors, by no means does it explain them all. Part of the reason for the lack of a perfect fit to the model is that there are large individual differences in how subjects view the chances of the marginal events. For instance (surprisingly), some subjects did not consider it more likely that "60 Minutes" would be the most watched television program of the 1984–1985 season than that the average conventional home mortgage interest rate would fall below 10%. As suggested by the protocols of Studies 2 and 3, another reason is that people simply use multiple methods in arriving at likelihood judgments.

It is highly adaptive that people have a broad repertoire of likelihood judgment procedures. This characteristic is reminiscent of the "equipotentiality" exhibited by neurological systems, whereby if one part of the brain is destroyed, other parts will sometimes assume functions ordinarily performed by the damaged part (Lashley, 1950). Thus, when judgment problems arise, we will seldom be caught "procedureless," as it were. Nevertheless, a fundamentally important question is raised by this wealth of procedures: How is a particular procedure assigned to a given judgment task? We have seen from Study 1 that most subjects appear to know the conjunction rule. So why do they not routinely apply it and thereby avoid making conjunction errors?

It will be some time before a full understanding of judgment procedure assignment principles is achieved. However, one candidate principle is that the characteristics of the judgment problem will significantly influence which procedure is called forth from the subject's repertoire. This principle seems to be the one implied by Tversky and Kahneman (1983) when they proposed (p. 310) that "a judgment of probability or frequency is commonly biased toward the natural assessment that the problem evokes." In Study 1 the Urn problem was correctly solved much more often than the naturalistic problems. In Studies 2 and 3 it was found that substantive reasoning was an especially common means by which subjects ordered the events they considered. These results suggest that, in realistic problems with interesting content, attention is so heavily consumed by the content that it simply does not occur to the person to apply well-understood principles such as extension.

APPENDIX

Table A1 describes the marginal events A and B for the judgment problems subjects considered in Studies 2–4. The problems are organized into Ensembles according to their subject matter. The Processes

TABLE A1
JUDGMENT PROBLEM FEATURES. STUDIES 2–4

Ensemble	Processes	Study	Chances	Event A	Event B
McEnroe	C[a]	4	LL[b]	McEnroe will win the match	McEnroe will win the first set
		2/3	LU	McEnroe will win the match	McEnroe will lose the first set
		4	UU	McEnroe will lose the match	McEnroe will lose the first set
Football/ Mideast	U	4	LL	The University of Michigan will win at least 4 of its 11 games during the 1984 football season	Syria and Israel will not sign a peace treaty with each other this year
		2/3	LU	The University of Michigan will win at least 4 of its 11 games during the 1984 football season	Syria and Israel will sign a peace treaty with each other this year
		4	UU	The University of Michigan will win fewer than 4 of its 11 games during the 1984 football season	Syria and Israel will sign a peace treaty with each other this year
Reagan	C	2/3	LU	President Reagan will provide for increased military spending	Reagan will provide federal support for unwed mothers
		4	UU	President Reagan will provide for decreased military spending	Reagan will provide federal support for unwed mothers

TABLE A1—*Continued*

Ensemble	Processes	Study	Chances	Event A	Event B
"60 Minutes"/ mortgages	U	4	LL	"60 Minutes" will be the most watched television program of the 1984–85 season	The average conventional home mortgage interest rate will remain above 10%
		3	LU	"60 Minutes" will be the most watched television program of the 1984–85 season	The average conventional home mortgage interest rate will fall below 10%
USSR	C	4	LL	The Soviet Union will keep its troops in Afghanistan	The Soviet Union will maintain its intermediate range missiles aimed at Western Europe
		2/3	LU	The Soviet Union will keep its troops in Afghanistan	The Soviet Union will dismantle its intermediate range missiles aimed at Western Europe
Weather/ AIDS	U	2	LU	It will be warmer on May 15 than on March 30	An effective treatment for AIDS (Acquired Immune Deficiency Syndrome) will be found within the year
		3	LU	It will be warmer on September 15 than on October 30	An effective treatment for AIDS (Acquired Immune Deficiency Syndrome) will be found within the year
		4	UU	It will be warmer on October 30 than on September 15	An effective treatment for AIDS (Acquired Immune Deficiency Syndrome) will be found within the year
Congress/ Derek	U	2	LU	Congress will avoid raising taxes in this election year	Bo Derek will be nominated for an Academy Award for her performance in her latest film
Election/ "Breakin' "	U	4	LL	Ronald Reagan will carry the South in the November Presidential election	The movie "Breakin' " will not be nominated for an Academy Award as "Best Picture"
		4	UU	Walter Mondale will carry the South in the November Presidential election	The movie "Breakin' " will be nominated for an Academy Award as "Best Picture"
Cars/ Tigers	U	4	LL	U.S. auto companies will raise their average car prices by at least $25	The Detroit Tigers will end the season with a better record than the Cleveland Indians
		4	UU	U.S. auto companies will raise their average car prices by less than $25	The Detroit Tigers will end the season with no better record than the Cleveland Indians

[a] C—common processes; U—unrelated processes.
[b] LL—event A and event B likely; LU—event A likely, event B unlikely; UU—event A and event B unlikely.

column indicates whether events A and B are generated by either common (C) or unrelated (U) processes. The Study column indicates the study in which a particular problem was presented. The Chances column notes whether each event was expected to be considered highly likely (L) or unlikely (U) by the typical subject. The expectation for event A is listed first, that for event B second.

Unless instructed otherwise, the subject was asked to evaluate the probabilities that various events considered in a given problem would occur within the next year, "between today and a year from today." The tennis match referred to in the McEnroe ensemble was McEnroe's first-round singles match in the 1984 U.S. Open tournament.

REFERENCES

Anderson, N. H. (1965). Averaging versus adding as a stimulus-combination rule in impression formation. *Journal of Experimental Psychology*, **70**, 394–400.

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Cochran, W. G. (1950). The comparison of percentages in matched samples. *Biometrika*, **37**, 256–266.

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

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, **87**, 215–251.

Gilhooly, K. J., & Falconer, W. A. (1974). Concrete and abstract terms and relations in testing a rule. *Quarterly Journal of Experimental Psychology*, **26**, 355–359.

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

Lashley, K. S. (1950). In search of the engram. In *Society of Experimental Biology Symposium No. 4: Physiological Mechanisms in Animal Behaviour* (pp. 454–482). Cambridge, England: Cambridge Univ. Press.

Markus, H., & Zajonc, R. (in press). The cognitive perspective in social psychology. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed.). Reading, MA: Addison–Wesley.

Morier, D. M., & Borgida, E. (1984). The conjunction fallacy: A task specific phenomenon? *Personality and Social Psychology Bulletin*, **10**, 243–252.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, **84**, 231–259.

Pitz, G. F., & Sachs, N. J. (1984). Judgment and decision: Theory and application. *Annual Review of Psychology*, **35**, 139–163.

Rapoport, A., & Wallsten, T. S. (1972). Individual decision behavior. *Annual Review of Psychology*, **23**, 131–175.

Shanteau, J. (1975). Averaging versus multiplying combination rules of inference judgment. *Acta Psychologica*, **39**, 83–89.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology*, **28**, 1–39.

Troutman, C. M., & Shanteau, J. (1977). Inferences based on nondiagnostic information. *Organizational Behavior and Human Performance*, **19**, 43–55.

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

Tversky, A., & Kahneman, D. (1982). Judgments of and by representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgments under uncertainty: Heuristics and biases* (pp. 84–98). New York: Cambridge Univ. Press.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, **90**, 293–315.

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

RECEIVED: January 2, 1985