BEHAVIORAL DECISION THEORY [1]

Paul Slovic, Baruch Fischhoff,

and Sarah Lichtenstein[2]

Oregon Research Institute

Annual Review of Psychology, 1977, in press.

Headings for Behavioral Decision Theory

DESCRIPTIVE RESEARCH

Probabilistic Judgment

MODEL-BASED PARADIGMS

Cascaded inference

HEURISTICS AND BIASES

Judgment by representativeness

Judgment by availability

Judgment by adjustment

Related work

Overconfidence

Descriptive theories

Choice

ELIMINATION BY ASPECTS

PROCESS DESCRIPTION

SCRIPT PROCESSING

CONSUMER CHOICE

Models of Risky Choice

Regression Approaches

INTEGRATION THEORY

POLICY CAPTURING

MULTIPLE CUE PROBABILITY LEARNING

Dynamic Decision Making

Are Important Decisions Biased?

EXPERTS IN THE LABORATORY

OUT IN THE FIELD

THE ULTIMATE TEST

Behavioral decision theory has two interrelated facets, normative and descriptive. The normative theory is concerned with prescribing courses of action that conform most closely to the decision maker's beliefs and values. Describing these beliefs and values and the manner in which individuals incorporate them into their decisions is the aim of descriptive decision theory.

This review is organized around these two facets. The first section deals with descriptive studies of judgment, inference and choice; the second section discusses the development of decision-aiding techniques.

As we reviewed the literature, several trends caught our attention. One is that decision making is being studied by researchers from an increasingly diverse set of disciplines, including medicine, economics, education, political science, geography, engineering, marketing, and management science, as well as psychology. Nevertheless, the importance

---

[1]This is the fourth survey of this topic to appear in the Annual Review. Its predecessors were articles by Edwards (78), Becker & McClintock (24), and Rapoport & Wallsten (226). The present review covers publications appearing between Jan. 1, 1971, and Dec. 31, 1975, with occasional exceptions.

[2]Support for this review was provided by the Advanced Research Projects Agency of the Department of Defense (ARPA Order No. 2449) and was monitored by the Office of Naval Research under Contract No. N00014-76-0074 (ARPA Order No. 3052) under subcontract from Decisions and Designs, Inc.

We wish to thank Barbara Combs, Robyn Dawes, Lewis R. Goldberg and Jerry LaCava for their comments on an early draft of the manuscript.

Nancy Collins and Peggy Roecker have earned our gratitude and respect for handling an arduous secretarial job with competence and good humor.

of psychological concepts is increasing, in both the normative and descriptive work. Whereas past descriptive studies consisted mainly of rather superficial comparisons between actual behavior and normative models, research now focuses on the psychological underpinnings of observed behavior. Likewise, the prescriptive enterprise is being psychologized by challenges to the acceptability of the fundamental axioms of utility theory (140, 188, 256).

Second, increasing effort is being devoted to the development of practical methods for helping people cope with uncertainty. Here, psychological research provides guidance about how to elicit the judgments needed for decision-aiding techniques.

Third, the field is growing rapidly, as evidenced by the numerous reviews and bibliographies produced during the past five years. Slovic & Lichtenstein (254) reviewed the literature on Bayesian and regression approaches to studying information processing in decision making and judgment; Dillon (73) covered utility theory with a view towards its application in agricultural contexts; MacCrimmon (187) examined work in management decision making; Shulman & Elstein (247) discussed the implications of judgment and decision making research for teachers; Nickerson & Feehrer (209) searched for studies relevant to the training of decision makers (since there aren't many, they settled for a general review); Vlek & Wagenaar (292) surveyed the entire field and Kozielecki (157) and Lee (165) have provided its first textbooks.

A selective and annotated bibliography on Behavioral Decision Theory has been compiled by Barron (18). Kusyszyn (161, 162) has provided bibliographies covering the psychology of gambling, risk-taking, and subjective probability. Houle (124) has accumulated a massive bibliography on Bayesian statistics and related behavioral work, which by 1975 included 106 specialized books, 1322 journal articles, and about 800 other publications. By the time you read this, Kleiter, Gachowetz & Huber (153) will have assembled the most complete bibliography ever in this field. They generously supplied us with more than 1000 relevant references, all produced between 1971 and 1975.

To ease cognitive strain (and stay within sight of our page allotment), we have focused on psychological aspects of individual judgment and decision making. Thus, we omit group and organizational decision making, Bayesian statistics, and much of the work on the axiomatic formulations of decision theory. Game theory is reviewed elsewhere in this volume. Even with this narrow focus, we have had to limit our coverage severely, concentrating on those references to which our prejudices have led us.

DESCRIPTIVE RESEARCH

Probabilistic Judgment

Because of the importance of probabilistic reasoning to decision making, considerable effort has been devoted to studying how people perceive, process and evaluate the probabilities of uncertain events. Early research on "intuitive statistics" led Peterson & Beach (218) to an optimistic conclusion:

> . . . man gambles well. He survives and prospers while using . . . fallible information to infer the states of his uncertain environment and to predict future events (p. 29).
>
> Experiments that have compared human inferences with those of statistical man show that the normative model provides a good first approximation for a psychological theory of inference. Inferences made by subjects are influenced by appropriate variables in appropriate directions (pp. 42-43).

MODEL-BASED PARADIGMS One result of this high regard for our intellectual capability has been a reliance on normative models in descriptive research. Thus, Barclay, Beach & Braithwaite (15) proposed beginning with a normative model and adjusting its form or parameters to produce a descriptive model. This approach is best exemplified by the study of conservatism--the tendency, when integrating probabilistic information, to produce posterior probabilities smaller than those specified by Bayes' theorem. In 1971, conservatism was identified as the primary finding of Bayesian information integration research (254). Reports of the phenomenon have continued to appear, in tasks involving normally distributed populations (75, 290, 305), and in that old favorite, the binomial (bookbag and poker chip) task (3, 196). Even filling the bookbags with male and female Polish surnames fails to lessen the effect (262). Donnell & DuCharme's (75) subjects became optimal when told the normative response, but when the task changed, their learning failed to generalize. As the next section shows, conservatism occurs only in certain kinds of inference tasks. In a variety of other settings, people's inferences are too extreme.

<u>Cascaded inference</u> Real-life problems often have several stages, with inferences at each stage relying on data which are themselves inferences from unreliable observations or reports. For example, a physician who uses the condition of the patient's lungs as a cue for diagnosis must first infer that condition from unreliable data (e.g., the sound of a thumped chest). Several normative models for such cascaded or multi-stage inference tasks have been developed in recent years (217, 238). Schum (239) has shown the relevance of cascaded inference models to the judicial problem of witness credibility and the probative value of witness testimony.

Descriptive studies of cascaded inference, comparing subjects' responses in the laboratory with a normative model, have consistently shown a result just the opposite of conservatism: subjects' posterior probabilities are more extreme than those prescribed by the model (100, 217, 266). The extremity of subjects' responses has been traced to their use of a simple, but inappropriate, "best-guess" strategy (103, 137, 257, 266), which is insensitive to data unreliability.

HEURISTICS AND BIASES

In these recent studies of conservatism and cascaded inference, one can see an increasing skepticism about the normative model's ability to fulfill its descriptive role, and the view of humans as good intuitive statisticians is no longer paramount. A psychological Rip van Winkle who dozed off after reading Peterson & Beach (218) and roused himself only recently would be startled by the widespread change of attitude exemplified by statements such as "In his evaluation of evidence, man is apparently not a conservative Bayesian: he is not Bayesian at all" (138, p. 450), or ". . . man's cognitive capacities are not adequate for the tasks which confront him" (114, p. 4), or ". . . people systematically violate the principles of rational decision making when judging probabilities, making predictions, or otherwise attempting to cope with probabilistic tasks" (252, p.   ).

Van Winkle would be further surprised to see Hammond (114) and Dawes (69) putting information-processing deficiencies on a par with motivational conflicts as causes of the ills that plague humanity, and to see financial analysts, accountants, geographers, statisticians and others being briefed on the implications of these intellectual shortcomings (14, 121a, 248, 249, 253, 282).

In 1971, when reviewing the literature on probabilistic inference, Slovic & Lichtenstein (254) found only a handful of studies that looked at subjects' information-processing heuristics. Since then, rather than simply comparing behavior with normative models, almost every descriptive study of probabilistic thinking has attempted to determine how the underlying cognitive processes are molded by the interaction between the demands of the task and the limitations of the thinker.

Much of the impetus for this change can be attributed to Tversky & Kahneman's (138, 139, 284, 285, 286) demonstrations of three judgmental heuristics, representativeness, availability and anchoring, which determine probabilistic judgments in a variety of tasks. Although always efficient, and at times valid, these heuristics can lead to biases that are large, persistent, and serious in their implications for decision making.

<u>Judgment by representativeness</u> What is the probability that object B belongs to class A? Or, what is the probability that process A will generate event B? Kahneman & Tversky (138) hypothesized that people answer such questions by examining the essential features of A and of B and assessing the degree of similarity between them, the degree to which B is "representative" of A. When B is very similar to A, as when an outcome is highly representative of the process from which it originates, then its probability is judged to be high.

Several lines of evidence support this hypothesis. Tversky & Kahneman (284) demonstrated a belief in what they called "the law of small numbers" whereby even small samples are viewed as highly representative of the populations from which they are drawn. This belief led their subjects, research psychologists, to underestimate the error and unreliability inherent in small samples of data. Kahneman & Tversky (138) showed that both subjective sampling distributions and posterior probability estimates were insensitive to sample size, a normatively important but psychologically non-representative factor. In a subsequent paper, Kahneman & Tversky (139) demonstrated that people's intuitive predictions violate normative principles in ways that can be attributed to representativeness biases. For one, representativeness causes prior probabilities to be neglected. For another, predictions tend not to be properly regressive, being insensitive to considerations of data reliability.

Judgment by availability Other judgmental biases are due to use of the "availability" heuristic (285) whereby an event is judged likely or frequent if it is easy to imagine or recall relevant instances. In life, instances of frequent events are typically easier to recall than instances of less frequent events, and likely occurrences are usually easier to imagine than unlikely ones. Thus, availability is often a valid cue for the assessment of frequency and probability. However, since availability is also affected by subtle factors unrelated to likelihood, such as familiarity, recency, and emotional saliency, reliance on it may result in systematic biases.

Judgment by adjustment Another error-prone heuristic is "anchoring and adjustment." With this process, a natural starting point or anchor is used as a first approximation to the judgment. The anchor is then adjusted to accommodate the implications of additional information. Typically, the adjustment is imprecise and insufficient (248). Tversky & Kahneman (286) showed how anchoring and adjustment could cause the overly narrow confidence intervals found by many investigators (175) and the tendency to misjudge the probability of conjunctive and disjunctive events (16, 57, 317).

Related work Numerous studies have replicated and extended the Kahneman & Tversky studies, and others have independently arrived at similar conclusions. The representativeness heuristic has received the most attention. Wise & Mockovak (310), Bar-Hillel (17), and Teigen (278, 279) have documented the importance of similarity structures in probability judgment. Like Kahneman & Tversky (138), Marks & Clarkson (191, 192) and Svenson (271) observed that subjects' posterior probabilities in binomial bookbag and poker chip tasks were predominantly influenced by the most representative aspect of the sample, the proportion of red chips. Contrary to the normative model, population proportion and sample size were relatively unimportant. Leon & Anderson (166) did find an influence of these two characteristics and, as a result, claimed that Kahneman & Tversky's subjects must have misunderstood the task. Ward (302), however, argued that the conflicting results were most likely due to differences in the tasks, rather than to misinterpretation of instructions. Hammerton (113), Lyon & Slovic (184), Nisbett & Borgida (210), and Borgida & Nisbett[3] have replicated Kahneman & Tversky's finding that subjects neglect population base rates when judging the probability that an individual belongs to a given category. Nisbett & Borgida argued that this neglect stems in part from the abstract, pallid, statistical character of base-rate information. They found that concrete, case-specific information, even from a sample of one, may have much greater importance, a rather dramatic

---

[3] Borgida, E. & Nisbett, R. E. Abstract vs. concrete information: The senses engulf the mind, unpublished, University of Michigan, 1976.

illustration of the law of small numbers. Additional evidence for representativeness comes from studies by Brickman & Pierce (45), Holzworth & Doherty (123), Bauer (20, 21) and Lichtenstein, Earle & Slovic (173).

Availability and anchoring have been studied less often. Evidence of availability bias has been found by Borgida & Nisbett[3] and Slovic, Fischhoff & Lichtenstein (252). Anchoring has been hypothesized to account for the effects of response mode upon bet preferences (176, 177) and it has been proposed as a method that people use to reduce strain when making ratio judgments (106). Pitz (219) gave the anchoring heuristic a key role in his model describing how people create subjective probability distributions for imperfectly known (uncertain) quantities.

Overconfidence The evidence presented above suggests that the heuristic selected, the way it is employed and the accuracy of the judgment it produces are all highly problem-specific; they may even vary with different representations of the same problem. Indeed, heuristics may be faulted as a general theory of judgment because of the difficulty of knowing which will be applied in any particular instance.

There is, however, one fairly valid generalization that may be derived from this literature. Except for some Bayesian inference tasks, people tend to be overconfident in their judgments. This may be seen in their non-regressive predictions (139), in their disregard for the extent of the data base upon which their judgments rest (138), or its reliability (217), and in the miscalibration of their probabilities for discrete and continuous propositions (175). Howell (128) has repeatedly shown that people overestimate their own abilities on tasks requiring skill (e.g., throwing darts). Langer (163) dubbed this effect "the illusion of control" and demonstrated that it can be induced by introducing skill factors (such as competition and choice) into chance situations.

In a task that had people estimate the odds that they had been able to select the correct answer to general knowledge questions, Slovic, Fischhoff & Lichtenstein (251) found that wrong answers were often given with certainty. Furthermore, subjects had sufficient faith in their odds that they were willing to participate in a gambling game that punished them severely for their overconfidence.

How do we maintain this overconfidence? One possibility is that the environment is often not structured to show our limits. Many decisions we make are quite insensitive to errors in estimating what we want (utilities) or what's going to happen (probabilities)--so that errors in estimation are hard to detect (294a). Sometimes we receive no feedback at all. Even when we do, we may distort its meaning to exaggerate our judgmental prowess, perhaps convincing ourselves that the outcome we got was what we really wanted. Langer & Roth (164) found that subjects who experienced initial successes in a repetitive task overremembered their own past successes. Fischhoff & Beyth (93) found that people asked to recall their own predictions about past events remembered having assigned higher probabilities to events that later occurred than was actually the case. Fischhoff (89) also found that people (a) overestimate the extent to which they would have been able to predict past events had they been asked to do so, and (b) exaggerate the extent to which others should have been able to predict past events. These hindsight biases are further evidence of overconfidence for they show that people have inordinately high opinions of their own predictive abilities.

<u>Descriptive theories</u> Most of the research on heuristics and biases can be considered pre-theoretical. It has documented the descriptive shortcomings of the normative model and produced concepts such as representativeness and anchoring that may serve as the basis for new descriptive theories. Although theory development has been limited thus far, efforts by Wallsten (300, 301) and Shanteau (243, 244) to produce descriptive algebraic models are noteworthy. Shanteau's approach is based upon the averaging model of Anderson's integration theory (7). Wallsten's model, formulated and tested within the framework of conjoint measurement, assumes that limited capacity causes people to process dimensions of information sequentially and weight them differentially, according to their salience.

Choice

In their introduction to two volumes on contemporary developments in methematical psychology, Krantz, et al (159) explained their exclusion of the entire area of preferential choice as follows:

> There is no lack whatever of technically excellent papers in this area but they give no sense of any real cumulation of knowledge. What are established laws of preferential choice behavior? (Since three of the editors have worked in this area, our attitude may reflect some measure of our own frustration). (p. xii)

This sense of frustration is understandable when one reviews recent research on choice. The field is in a state of transition, moving away from the assumption that choice probability is expressable as a monotone function of the scale values or utilities of the alternatives. Present efforts are aimed at developing more detailed, molecular concepts, that describe choice in terms of information-processing phenomena. Researchers appear to be searching for heuristics or modes of processing information that are common to a wide domain of subjects and choice problems. However, they are finding that the nature of the task is a prime determinant of the observed behavior.

ELIMINATION BY ASPECTS One major new choice theory is Tversky's (280, 281) elimination-by-aspects (EBA) model. The model describes choice as a covert sequential elimination process. Alternatives are viewed as sets of aspects (e.g., cars described by price, model, color, etc.). At each stage in the choice process an aspect is selected with probability proportional to its importance; alternatives that are unsatisfactory on the selected aspect are eliminated. Tversky showed that the EBA model generalizes the models of Luce (183) and Restle (228) while avoiding some of the counter-examples to which these earlier models are susceptible. Searching for even broader applicability, Corbin & Marley (62) proposed a random utility model that includes the EBA model as a special case. Other models built around the concept of successive elimination of alternatives have been developed by Hogarth (121, 122) and Pollay (220).

PROCESS DESCRIPTIONS Most recent empirical research has been concerned with describing the decision maker's methods for processing information before choosing. Whereas earlier work focused on external products (e.g., choice proportions and rankings) and used rather simple methods, process-descriptive studies must employ more complex procedures for collecting and analyzing data. Thus, we find a return to introspective methods (28, 199, 272) in which subjects are asked to think aloud as they choose among various multiattribute alternatives. Bettman & Jacoby (31) and Payne (214) supplemented the think-aloud procedure by requiring subjects to seek information from envelopes on an "information board." Russo & Rosen (231) used eye-movement data conjointly with verbal protocols. One goal of these studies is to represent the choice process graphically as a tree or network (discrimination net) of successive decisions. Swinth, Gaumnitz & Rodriguez (275) developed a method of controlled introspection that enables subjects to build and validate their own discrimination nets. Bettman (27) showed how to describe such nets via graph-theoretical concepts. Uneasy about the subjectivity of introspective techniques, Hogarth (121) used an ingenious blend of theory and empiricism to develop a computer algorithm that builds the tree without recourse to subjective inputs.

Can introspective methods be trusted? Nisbett & Wilson[4] reopened an old debate by arguing that people lack awareness of the factors that affect their judgments. After documenting this claim with results from six experiments, they concluded that "Investigators who are inclined to place themselves at the mercy of such [introspective] reports . . . would be better advised to remain in the armchair" (p. 35). While important, this criticism may be overstated. Students of choice have in many instances validated their introspective reports against theoretical predictions (199) and data from other sources (214, Footnote 5).

What do these methodologies tell us about choice? First they indicate that subjects use many rules and strategies enroute to a decision. These

---

[4]Nisbett, R. E. & Wilson, T. D. Awareness of factors influencing one's own evaluations, judgments, and behavior, unpublished, University of Michigan, 1976.

include conjunctive, disjunctive, lexicographic and compensatory rules and the principle of dominance (274). A typical choice may involve several stages, utilizing different rules at different junctures. Early in the process, subjects tend to compare a number of alternatives on the same attribute and use conjunctive rules to reject some alternatives from further consideration (26, 214, 245, 272). Later they appear to employ compensatory weighting of advantages and disadvantages on the reduced set of alternatives (214). Features of the task that complicate the decision, such as incomplete data, incommensurable data dimensions, information overload, time pressures and many alternatives seem to encourage strain-reducing, noncompensatory strategies (214, 255, 313, 314). Svenson (272) and Russo & Rosen (231) found subjects reducing memory load by comparing two alternatives at a time and retaining only the better one for later comparisons. Russo & Dosher[5] observed simple strategies, such as counting the number of dimensions favoring each alternative or ignoring small differences between alternatives on a particular dimension. In some instances, these strategies led to sub-optimal choices.

In general, people appear to prefer strategies that are easy to justify and don't involve reliance on relative weights, trade-off functions or other numerical computations. One implication of this was noted by Slovic (250), whose subjects were forced to choose among pairs of alternatives that were equal in value for them. Rather than choose randomly, subjects consistently followed the easy and defensible strategy of selecting the

---

[5]Russo, J. E. & Dosher, B. A. Dimensional Evaluation: A heuristic for binary choice, unpublished, University of California, Santa Barbara, 1975.

alternative that was superior on the more important dimension.

SCRIPT PROCESSING Abelson's (1) new approach to explaining decisions warrants further study. It is based on the concept of a "cognitive script," which is a coherent sequence of events expected by the individual on the basis of prior learning or experience. When faced with a decision, individuals are hypothesized to bring relevant scripts into play. For example, Candidate Y's application for graduate school may be rejected because Y reminds the decision maker of Candidate X who was accepted and failed miserably. Another script might assimilate the candidate into a category (He's one of those shy types who does well in courses, but doesn't have enough initiative in research). Script theory, though still in a highly speculative stage, suggests a type of explanation for choice that has thus far been overlooked.

CONSUMER CHOICE Much research on choice has been done within the domain of consumer psychology. Comprehensive reviews of this research have been provided by Jacoby (134, 135). Although some of this work is application of multiple regression, conjoint measurement, and analysis of variance to describe consumers' values (30, 107, 312), many other studies have investigated basic psychological questions. For example, one major issue has been the effect of amount and display of information on the optimality of choice. Jacoby and his colleagues have argued that more information is not necessarily helpful, as it can overload consumers and lead them to select sub-optimal products. Russo, Krieser & Miyashita (230) observed that subjects had great difficulty finding the most economical product among an array of different prices and packages. Even unit prices, which do the arithmetic for the consumer, had little effect on buyer behavior when posted on the shelf below each product. However, when prices per unit were listed in order from high to low cost, shoppers began to buy less expensive products.

Models of Risky Choice

Decision making under conditions of risk has been studied extensively. This is probably due to the availability of (a) an appealing research paradigm, choices among gambles, and (b) a dominant normative theory, the subjectively expected utility (SEU) model, against which behavior can be compared. The SEU model assumes that people behave as though they maximized the sum of the products of utility and probability.

Early studies of the model's descriptive adequacy produced conflicting results. Situational and task parameters were found to have strong effects, leading Rapoport & Wallsten (226) to observe that a researcher might accept SEU with one set of bets and reject it with another, differently structured set. Proponents of the SEU model point out that it gives a good global fit to choice data, particularly for simple gambles.[6] In addition, certain assumptions of the model, like the independent (multiplicative) combination of probabilities and payoffs, have been verified for simple gambles (244, 299).

However, during the past five years, the proponents of SEU have been greatly outnumbered by its critics. Coombs (60) has argued that risky choice is determined not by SEU, but by a compromise between maximization

---

[6] Goodman, B., Saltzman, M., Edwards, W. & Krantz, D. Prediction of bids for two-outcome gambles in a casino setting, unpublished, 1976.

of expected value (EV) and optimization of risk. He proposed an alternative to SEU, "portfolio theory," in which risk preferences play a central role. That role is illustrated in a study by Coombs & Huang (61) in which a gamble, B, was constructed as a probability mixture of two other gambles, A and C. Many subjects preferred gamble B (with its intermediate risk level) to gambles A and C, thus violating a fundamental axiom of SEU theory.

Zagorski (318) demonstrated a result that appears to violate SEU and many other algebraic models as well. Zagorski's subjects were shown pairs of gambles (A, B) and were asked to judge the amount of money (A-B) that would induce them to trade the better gamble (A) for the worse gamble (B). He demonstrated that one can construct quadruples of gambles A, B, C and D such that

$$(A-B) + (B-C) \neq (A-D) + (D-C) .$$

In other words, path independence is violated. The difference between gambles A and C depends on whether the intermediate gamble is B or D.

A favorite approach of SEU critics is to develop counterexamples to the fundamental axioms of the theory. The paradoxes of Allais (4) and Ellsberg (85) are two of the most famous, both designed to invalidate Savage's (232) independence principle. Until recently, few theorists were convinced. MacCrimmon (185) showed that business executives who violated various axioms could easily be led, via discussion, to see the error of their ways. However, Slovic & Tversky (256) challenged MacCrimmon's discussion procedure on the grounds that it pressured the subjects to accept the axioms. They presented subjects with arguments for and against the independence axiom and found persistent violations, even after the axiom was presented in a clear and presumably compelling fashion. Moskowitz (200) used a variety of problem representations (matrix formats, trees, and verbal presentations) to clarify the principle and maximize its acceptability, yet still found that the independence axiom was rejected. Even MacCrimmon's faith in many of the key axioms has been shaken by recent data (see 188), leading him to suggest that reevaluation of the theory is in order.

Kahneman & Tversky (140, 283) attempted this sort of reevaluation, presenting evidence for two pervasive violations of SEU theory. One, the "certainty effect," causes consequences that are obtained with certainty to be valued more than uncertain consequences. The Allais paradox may be due to this effect. The second, labeled the "reference effect," leads people to evaluate alternatives relative to a reference point corresponding to their status quo, adaptation level, or expectation. By altering the reference point, formally equivalent versions of the same decision problem may elicit different preferences. These effects pose serious problems for the normative theory and its application.

Payne (213) proposed replacing the SEU model with information processing theories that describe how probabilities and payoffs are integrated into decisions. He presented a "contingent process" model to describe the sequential processes involved in choice among gambles. For support, he cited a number of display and response-mode effects that are due to processing difficulties (176, 177, 179, 215). Kozielecki's (158) discussion of the internal representation of risky tasks carried a similar message.

Kunreuther (160) has argued that utility theory would be of little value to a policy maker trying to predict how people would respond to various flood or earthquake insurance programs. First, the theory makes predictions that are not borne out by actual behavior--for example, that people will prefer policies with high deductibles or that subsidizing premiums will increase insurance purchasing. Second, it gives no guidance about the social, situational and cognitive factors that are likely to influence insurance purchase. Like Payne, Kunreuther called for an alternative theory, founded on the psychology of human information processing, and presented a model of his own to support his case.

Readers interested in additional attacks on the staggering SEU model should consult Barron & MacKenzie (19), Davenport & Middleton (66), Fryback, Goodman & Edwards (99), Ronen (229), and Svenson (273).

<u>Regression Approaches</u>

The regression paradigm uses analysis of variance, conjoint measurement and multiple regression techniques to develop algebraic models that describe the method by which individuals weight and combine information.

INTEGRATION THEORY Working within the framework of "information integration theory," Anderson and his colleagues have shown that simple algebraic models describe information use quite well in an impressive variety of judgmental, decision making, attitudinal, and perceptual tasks (6, 7). These models typically have revealed stimulus averaging, although some subtracting and multiplying has been observed. Particularly relevant to decision making are studies of risk taking and inference (244), configurality in clinical judgment (5), intuitive statistics (167, 168), preference for bus transportation (210a), and judgment in stud poker (181). There is no doubt that algebraic models derived from Anderson's techniques provide good surface descriptions of judgmental processes. However, as

Graesser & Anderson (106) have observed, establishment of an algebraic model is only the first step towards disclosing the underlying cognitive mechanisms, which may be rather different from the surface form of the model.

POLICY CAPTURING Another form of the regression paradigm uses correlational statistics to provide judgmental models in realistic settings. The most systematic development of these procedures has been made by Hammond and his colleagues (117) within "social judgment theory." This theory assumes that most judgments depend upon a mode of thought that is quasi-rational, that is, a synthesis of analytic and intuitive processes. The elements of quasi-rational thought are cues (attributes), their weights, and their functional relationships (linear and non-linear) to both the environment and the judge's responses. Brusnwik's lens model and multiple regression analysis are used to derive equations representing the judge's cue utilization policy. Judgmental performance is analyzed into knowledge and "cognitive control," the latter being the ability to employ one's knowledge consistently (118).

By 1971, it was evident that linear models could describe college students' cue-weighting policies in a wide variety of laboratory tasks (254). During the past five years, such models have been used with similar success to analyze complex real-world judgments. Judges in these studies have included business managers (119, 193, 201, 202), graduate admissions committees (68, 237), auditors, accountants, and loan officers (13, 172, 315), military officers (277), literary critics (84), and trout hatchery employees (182), as they attempted to predict business failures and stock market performance, select graduate students, plan work force and production schedules, evaluate accounting procedures, Air Force cadets, and theatrical plays, and recommend trout streams. Even U.S. senators have been modeled and their roll-call votes predicted (298). As in the laboratory studies, linear equations have accounted for most of the predictable variance in these complex judgments. The coefficients of these equations have provided useful descriptions of the judges' cue-weighting policies and have pinpointed the sources of inter-judge disagreement and non-optimal cue use.

While policies were being captured in the field, other researchers were deepening our understanding of the models. Dawes & Corrigan (70) observed that linear models have typically been applied in situations in which (a) the predictor variables are monotonically related to the criterion (or can be easily rescaled to be monotonic), and (b) there is error in the independent and dependent variables. They demonstrated that these conditions insure good fits by linear models, regardless of whether the weights in such models are optimal. Thus the linearity observed in judges' behaviors may be reflecting only a characteristic of linear models, not a characteristic of human judgment.

In other work, theoretical and methodological refinements of the lens model have been developed by Castellan (52, 53) and Stenson (267). Cook (59) and Stewart & Carter (268) have worked towards developing interactive computer programs for policy capturing. Mertz & Doherty (195) and Brehmer (37) examined the influence of various task characteristics on the configurality and consistency of policies. Miller (197) demonstrated that improper cue labels could mislead judges despite the availability of adequate statistical information about cue validities. Lichtenstein, Earle & Slovic (173) and Birnbaum (32) showed that even though regression equations can be used to describe cue-combination policies, subjects often average cues, in violation of the additivity inherent in the equations.

in violation of the additivity inherent in the equations. Wiggins (306) discussed the problems of identifying and characterizing individual differences in judgmental policies and Ramanaiah & Goldberg (222) explored the stability and correlates of such differences. McCann, Miller & Moskowitz (193) examined the problems of capturing policies in particularly complex and dynamic tasks such as production planning.

MULTIPLE CUE PROBABILITY LEARNING Considerable effort has been invested in studying how people learn to make inferences from several probabilistic cues. Most of this work goes under the label "multiple-cue probability learning" (MCPL) and relies on the lens model for conceptual and analytic guidance. Typically, the cues are numerical and vary in their importance and in the form (linear or nonlinear) of their relationship to the criterion being judged. The criterion usually contains error, making perfect prediction impossible. Because these tasks embody the essential features of diagnostic inference, they are studied for their potential applied significance as well as their contributions to basic knowledge.

Slovic & Lichtenstein (254) reviewed MCPL studies published prior to 1971. They concluded that: (a) subjects can learn to use linear cues appropriately; (b) learning of nonlinear functions is slow, and especially difficult when subjects are not forewarned that relations may be nonlinear; (c) subjects are inconsistent, particularly when task predictability is low; (d) subjects fail to take proper account of cue intercorrelations; and (e) outcome feedback is not very helpful.

Research during the past half decade has confirmed and extended these conclusions. Difficulties people have in coping with intercorrelated cues have been documented in numerous studies (8, 9, 178, 236). Hammond and his colleagues (115) used the MCPL paradigm to analyze the effects of psychoactive drugs on cognition. They found that some drugs that are used to enhance emotional control interfered with learning and communication in ways that may be detrimental to therapy. Bjorkman (33) and Castellan (54) reviewed results from studies using nonmetric cues and criteria.

Other research has worked towards developing a theory to explain MCPL results in terms of erroneous intuitions about probabilistic tasks, the manner in which individuals acquire and test hypotheses, and their cognitive limitations. For example, Brehmer (38, 40, 41) has studied how subjects formulate and test hypotheses as they search for rules that will produce satisfactory inferences. Hypotheses about the functional rule relating cues and criterion appear to be sampled from a hierarchical set based on previous experience and dominated by the positive linear rule. Testing of hypotheses about rules shows inadequate appreciation of the probabilistic nature of the task. Subjects keep searching for deterministic rules that will account for the randomness in the task; since there are none, they change rules frequently (i.e., become inconsistent) and eventually resample rules they had previously discarded.

Even when subjects are informed of the correct rules, they have trouble applying them consistently (31, 42, 118). Nonlinear rules are particularly hard to apply. Brehmer, Hammond and their colleagues have thus conceptualized inference as a skill analogous to motor behavior: with both, we can know what we want to do without necessarily being able to do it.

Dynamic Decision Making

At the time of Rapoport & Wallsten's review, one active research area was dynamic decision making (DDM), the study of tasks in which "decisions are made sequentially in time; the task specifications may change over time, either independently or as a result of previous decisions; information available for later decisions may be contingent upon the outcomes of earlier decisions; and implications of any decision may reach into the future" (224, p. 345). The present half-decade began promisingly with Rapoport & Burkheimer's (225) explication of formal models for deferred decision making and the manner in which they might be utilized in psychological experiments. Shortly thereafter, Ebert (77) reported finding no difference between stochastic and deterministic versions of a task which Rapoport (223) earlier had found to differ. After that, relative silence.

Several possible reasons for this decline in interest come to mind. The mathematical sophistication of DDM may deter some researchers, as may the on-line computer and long start-up time often required. Furthermore, DDM models are so complex and require so many assumptions that the interpretation of experimental results is typically ambiguous--witness the morass of explanations facing Ebert (77) for why his experiment and Rapoport's produced different results. Kleiter (151) noted particular problems with creating cover stories that induce subjects to accept the assumptions underlying the model and with ascertaining that subjects understood the task. He also questioned "the metahypothesis that human behavior is optimal" (p 374), which limits psychological theories to variations on the optimal model, (e.g., using subjective probability estimates rather than "objective" relative frequencies or assuming a reduced planning horizon). In his own work, Kleiter (152) has assessed people's planning horizons and has used a non-normative variance-preference model to predict betting behavior in a multistage game (154). These predictions relied on the assumption that people were perfect Bayesian information processors.

A more active area of DDM research deals with sequential information purchasing or sampling. Levine & Samet (169) allowed subjects to purchase information from three fallible sources until they could decide which of eight possible targets was the object of an enemy advance. They found that information seeking decreased with conflicting and unreliable information, as did accuracy. On the other hand, Snapper & Peterson (259) reduced the diagnosticity of information and found people purchasing more. Their subjects appeared to be unresponsive to changes in information quality because of a policy of purchasing "intermediate" amounts of information.

Another sequential task that has attracted some attention is optional stopping: the decision maker must choose between accepting a currently available outcome versus sampling further outcomes that may be of greater or lesser worth. Although earlier research (see 225a) found that subjects performed well when options were generated by a random but stationary process, Brickman (44) found very poor performance with options that tended to increase or decrease in value. In particular, subjects persisted much longer in sampling options with a descending than with an ascending sequence. Brickman likened this behavior to "throwing good money after bad." His subjects' "take the money and run" strategy with ascending series was similar to that found by Corbin, Olson & Abbondanza (63). Their subjects seem to have called it quits as soon as an options appeared that was a good bit better than its predecessors. Ölander (212), too, described satisficing (rather than maximizing) principles that may guide subjects' decisions about searching further.

Are Important Decisions Biased?

A coherent picture emerges from research described so far. Because of limited information-processing capacity and ignorance of the rules for optimal information processing and decision making, people's judgments are subject to systematic biases. Can these results be generalized from the lab to the real world?

A number of critics are doubtful. Edwards (80) argued that experimenters, by denying subjects necessary tools and providing neither the time nor the guidance to find them, have exaggerated human intellectual limitations. Winkler & Murphy (309) criticized laboratory experiments for being overly simplified and too well structured when compared with the real-world situations they are meant to model. They suggested that people may perform poorly in the lab because of improper generalization from their real-world experiences. For example, because real world information tends to be redundant and unreliable, people may naturally devalue the reliable information provided in experiments, producing conservatism. In addition, experimental subjects may be poorly motivated and forced to deal with unfamiliar tasks and substantive areas, without adequate training--even in the meaning of the response mode (121a).

In rebuttal, one could argue that laboratory studies may show subjects at their best. Use of unfamiliar substantive topics may free them from preconceived notions that could prejudice their judgments. Provision of all information necessary for an optimal decision (and little else) is, as noted by Winkler & Murphy (309), a boon seldom offered by the real world. It may create demand characteristics forcing subjects toward optimal responses (90, 97, 302). An alternative rebuttal is that there are many real-life situations which are quite like the laboratory, forcing people to make a decision without the benefit of training and experience. People typically buy cars and houses and decide to marry and divorce under such circumstances, functioning as their own best approximation to experts.

Perhaps the best way to resolve this argument is to look at the evidence.

EXPERTS IN THE LABORATORY The robustness of biases is shown in formal experiments using experts as subjects. As examples: Tversky & Kahneman's (284) "law of small numbers" results were obtained with statistically savvy psychologists. Las Vegas casino patrons showed the same irrational reversals of preferences for gambles as did college students (176, 177). Bankers and stock market experts predicting closing prices for selected stocks showed substantial overconfidence and performed so poorly that they would have done better with a "know nothing" strategy (265). Lichtenstein & Fischhoff (174) found that the probability assessments of psychology graduate students were no better for questions within their area of expertise than for questions relating to general knowledge.

The "experts" in these studies were selected on the basis of what they knew about the subject area, not what they knew about judgment and decision making (i.e., they were substantive rather than normative experts). Can normative experts be created in the laboratory by proper training? The evidence is mixed, suggesting either that some biases are robust or that we have failed to understand the psychology of our subjects well enough to assist them.

OUT IN THE FIELD With the exception of some well-calibrated weather forecasters (described below), similar biases have been found in a variety of field studies. For example, Brown, Kahr & Peterson (49) observed overestimation in the probability assessments of military intelligence analysts. Kidd (149) found that engineers for the United Kingdom's Central Electricity Generating Board consistently underestimated repair time for inoperative units. Bond (34) observed suboptimal play among 53 blackjack players at four South Lake Tahoe casinos. "By wagering small bets in a sub-fair game, [these] blackjack gamblers practically guaranteed loss of their betting capital to the casinos" (p. 413). Flood plain residents misperceive the probability of floods in ways readily explained in terms of availability and representativeness (253). Surveying research published in psychological and educational journals, Cohen (56) and Brewer & Owen (43) found that investigators regularly design experiments with inadequate statistical power, reflecting a belief in the "law of small numbers" (284). Misinterpretation of regression toward the mean appears to be as endemic to some areas of psychology (101) as to Kahenman & Tversky's (139) subjects.

A major legal debate concerns the incarceration of individuals for being "dangerous." What little evidence there is regarding the validity of dangerousness judgments indicates substantial "over-prediction," incarceration of people who would not have misbehaved had they been set free (72, 242). Although this bias may reflect a greater aversion to freeing someone who causes trouble than to erring in the other direction, some observers have attributed it to judgmental problems such as failure to consider base rates, ignorance of the problems of predicting rare events, perception of non-existent correlations, and insensitivity to the reliability of evidence (198a).

Jurors appear to have great difficulty ignoring first impressions of the accused's personality, pretrial publicity, and other forms of inadmissible evidence (46, 270), tendencies which may represent both hindsight and anchoring biases (92). The vagaries of eyewitness testimony and witnesses' overconfidence in erroneous knowledge are quite well known (51, 180).

Zieve (319) has described at length the misinterpretation and abuse of laboratory test results by medical clinicians. Although some of these errors are due to ignorance, others reflect naive statistical reasoning. A classic case of the "law of small numbers" is Berkson, Magath & Hurn's (25) discovery that aspiring lab technicians were expected by their instructors to show greater accuracy in performing blood cell counts than was possible given sampling variation. These instructors would marvel that the best students (those who would not cheat) had the greatest difficulty in producing acceptable counts. In a phenomenological study of orthopedic surgeons, Knafl & Burkett (155) found a variety of simplifying heuristics, some of them in the form of general treatment philosophies (e.g., "don't cut unless you absolutely have to").

The immense decisions facing our society (e.g., nuclear power) have prompted the development of formal analytic techniques to replace traditional, error-prone,"seat of the pants" decision making. Fischhoff (91) reviewed a variety of cost-benefit analyses and risk assessments performed with these techniques and found them liable to omissions of important consequences reflecting availability biases. In case studies of policy analyses, Albert Wohlstetter (311) found that American intelligence analysts consistently underestimated Soviet missle strength, a bias possibly due to anchoring. Roberta Wohlstetter's (311a) study of American unpreparedness at Pearl Harbor found the U.S. Congress and military investigators guilty of hindsight bias in their judgment of the Pearl Harbor command staff's negligence.

Even if policy analyses are performed correctly, they still must be explained (sold?) to the public. In the area of natural hazard management, well-founded government policies have foundered because people don't perceive flood hazards the way policy makers expect them to (253). For example, the National Flood Insurance Program has had only limited success because the endangered won't buy the highly subsidized and normatively very attractive insurance offered them (160).

THE ULTIMATE TEST "If behavioral decision theory researchers are so smart, why aren't they rich?"

"They're not in business."

"Then why aren't people who are in business falling over themselves to utilize their results?"

Well, although psychological research has not swept the world's decision makers like wildfire, it has kindled some non-negligible interest. The concern weather forecasters and decision analysts have shown for research in probability assessment is described elsewhere in this review. The Department of Defense is developing sophisticated decision aids to relieve military commanders of the need to integrate information in their heads (148). U. S. intelligence analysts have shown interest in the use of Bayesian approaches for processing of intelligence information (79a, 147). Researchers in accounting (14, Footnote 7) have advocated considering information-processing limits in designing financial reports. The American College of Radiology has launched a massive "Efficacy Study" to see how radiologists use the probabilistic information from x-rays. Bettman (29), Armstrong, Kendall & Ross (11) and others have argued that legislation intended to provide consumers with necessary information (e.g., unit pricing, true interest

---

[7] Climo, T. A., Cash flow statements for investors, unpublished, University of Kent at Canterbury, 1975.

rates) must consider how those consumers do, in fact, process information.

DECISION AIDS

"What do you do for a living?"

"Study decision making."

"Then you can help me. I have some big decisions to make."

"Well, actually . . . "

That sinking feeling of inadequacy experienced by many of us doing psychological research in decision making is probably <u>not</u> felt by most experts in decision analysis, multiattribute utility theory or other decision aiding techniques. Proponents of these approaches have remedies for what ails you--techniques to help users make better decisions in any and all circumstances.

Most of these decision aids rely on the principle of divide and conquer. This "decomposition" approach is a constructive response to the problem of cognitive overload. The decision aid fractionates the total problem into a series of structurally-related parts, and the decision maker is asked to make subjective assessments for only the smallest components. Such assessments are presumably simpler and more manageable than assessing more global entities. Research showing that decomposition improves judgment has been reported by Armstrong, Denniston & Gordon (10), Gettys et al (104), and by Edwards and his colleagues (254, pp. 717-21).

Critics of the decomposition approach would argue that many of the aids require assessments of quantities the decision maker has never thought about, and that these apparently simple assessments may be psychologically more complex than the original decision. In some situations, people may really know what they want to do better than they know how to assess the inputs required for the decision aid.

Decision aids which do not rely on decomposition, but instead require the decision maker to state preferences among whole, nonfractionated alternatives, are here called "wholistic." The models in these aids are used to smooth or correct the wholistic judgments, and to partial them into components.

Since several of the decision aids rely on assessments of probability, we start this section with a review of probability elicitation techniques.

Assessing Probabilities

What's the best way to assess probabilities? Spetzler & Stäel von Holstein (260) have written an excellent description of how the Decision Analysis Group at Stanford Research Institute approaches this problem. They recommended (a) carefully structuring the problem with the client ("mental acrobatics should be minimized", p. 343), (b) minimizing biases that might affect the assessor, (c) using personal interviews rather than computer-interactive techniques with new clients, and (d) using several different elicitation methods, both direct and indirect. Their favorite elicitation technique is a reference bet involving a "probability wheel," a disk with two differently colored sectors whose relative size is adjustable. The assessor is offered two bets, each with the same payoff. One bet concerns the uncertain quantity (you win if next year's sales exceed $X); the other bet concerns the disk (you win if the pointer lands in the orange sector after the disk is spun). The relative size of the two sectors is varied until the assessor is indifferent between the two bets. The proportion of the disk which is orange is taken as the probability of the event stated in the other bet.

Despite the appeal of this method (it is formally justified within axiomatic models of subjective probability, does not require the assumption that the utility of money is linear with money, and requires no numerical response from the assessor), we have been unable to find any research on its use.

DISCRETE EVENTS Comparisons among several direct methods for assessing the probabilities of discrete events (probabilities vs. odds vs. log odds) have failed to identify one clearly preferable response mode (35, 73a, 105). Beach (22) found a mean within-subject correlation of only .49 between probabilities assessed directly and indirectly (via bids for bets). DuCharme & Donnell (76) found equally conservative inferences using odds, probabilities, and an indirect method similar in concept to, but more complicated than, the reference bet method discussed by Spetzler & Stäel von Holstein(260).

These studies focused on the assessment of middle-range probabilities; even less is known about assessing very large or very small probabilities. Slovic, Fischhoff & Lichtenstein (251) have shown that subjects grossly misuse odds of greater than 50:1. Selvidge (241) has made some common-sense suggestions for assessing very small probabilities. She advised first structuring and decomposing the problem, then ranking various unlikely events, and finally attaching numbers to those events with the help of reference events (like dying in various rare accidents).

Once you've assessed a probability, how good is it? When there is an agreed-upon "true probability"--as with bookbag and poker chip tasks--the assessed probability may be compared with the "truth." But more often, assessed probability may be compared with the "truth." But more often, the assessed probability states a degree of belief in some proposition, so that no criterion "true" probability value exists. One test of such probabilities is coherence, that is, do they abide by the axioms of probability? (290, 316). A second kind of validity, called calibration, may be examined if one collects a large number of assessments for which the truth of the associated propositions is known. For discrete propositions, calibration means that for every collection of propositions assigned the same numerical probability, the hit rate or proportion which actually are true should be equal to the assessed probability. The research on calibration has recently been extensively reviewed extensively (175), so only a summary of findings will be given here: (a) Experienced weather forecasters, when performing their customary tasks, are excellently calibrated. (b) Everybody else stinks. (c) People are overconfident except with very easy tasks.

UNCERTAIN QUANTITIES The most common technique for assessing probability density functions across uncertain quantities is the fractile method. An assessor who names a value of an uncertain quantity as its .25 fractile, for example, is saying that there is just a 25% chance that the true value will be smaller than that specified value. Stäel von Holstein (264) and Vlek (290) have studied the consistency between the fractile method and other elicitation methods. Stäel von Holstein found that even after four sessions most subjects were inconsistent. Vlek's subjects showed greater consistency.

Continuous probability density functions can also be tested for calibration. Assessors are calibrated when, over many such assessments, the proportion of true answers falling below a given fractile is equal to that fractile. The evidence on calibration (175) may be summarized as follows: (a) A strong and nearly universal bias exists: the assessed distributions are too tight, so that from 20% to 50% of the true values, instead of 2%, fall outside of the .01 to .99 range of the distributions; (b) Training improves performance.

SCORING RULES Scoring rules are functions which assign a score to an assessed probability (or a vector of probabilities) as a function of both the true outcome of the event being assessed and the size of the probability associated with the true outcome. Such rules are strictly proper if and only if the only strategy for maximizing one's expected score is to tell the truth--to state one's true belief without hedging. Usually the only rules considered are those which reward expertise: given that one tells the truth, the more one knows, the larger the score (an exception is Vlek's [291] fair betting game). Scoring rules have recently been discussed by Murphy & Winkler (205, 206) and by Shuford & Brown (50, 246).

Scoring rules may be used for three purposes. One use is as an indirect method for measuring probabilities. A list of bets is generated from the scoring rule. Each bet gives two numbers, how much the assessor wins if the event in question occurs and how much is lost if it does not. The assessor selects his or her preferred bet from the list; this choice implies a probability. Jensen & Peterson (136) and Seghers, Fryback & Goodman (240) found this method unsatisfactory; their subjects were apparently using other strategies rather than trying to maximize winnings.

Secondly, scoring rules may be used to educate assessors about probability assessments made with other methods. Several studies have used scoring rule feedback (246, 263, 308) without reporting whether it helped. Hoffman & Peterson (120) reported that subjects who received such feedback improved their scores on a subsequent task but Vlek (290) found no such improvement. Scoring rules are now widely used by weather forecasters, and this may be why they are so well calibrated (175). Murphy & Winkler (207) reported that a majority of 689 weather forecasters (a) described themselves as being uncomfortable thinking in probabilistic terms (though their job is to report probabilities and they do it well) and (b) rejected the idea that their forecasts can be properly evaluated by a single quantitative measure like scoring rule (though many had had experience with such feedback).

The third use for scoring rules is to evaluate assessors. When all assessors are working in the same situation, the assessor with the higher score is the better assessor. However, not all situations are equal; there is more uncertainty in forecasting rain in Chicago than in Oregon. Thus Oregon forecasters will earn higher scores simply because of where they work. Murphy (203) has shown that the Brier scoring rule (the one used in meteorology) may be partitioned into three additive components, measuring (a) the inherent uncertainty in the task, (b) the resolution of the assessor (i.e., the degree to which the assessor can successfully assign probabilities different from the overall hit rate), and (c) the assessor's calibration. None of the components is itself a proper scoring rule, but the difference between the total score and the inherent uncertainty component is proper, and this difference could be used to compare assessors in different situations (204).

The astute reader will note that the research does not provide an adequate answer to the question asked at the start of this section: What's the best way to assess probabilities? In addition, the research has yielded few theoretical ideas. Only Pitz (219) has speculated on the cognitive processes underlying probabiltiy assessment. Finally, although a few studies have noted that training improves performance in eliciting probabilities, a definitive long-range learning study is still needed.

Multiattribute Utility Theory

Suppose you must choose one object or course of action from a set. Each object or action is describable in terms of a number of dimensions or attributes of value to you, and the outcomes of your choice are certain. Then multiattribute utility theory (MAUT) prescribes that you compute, for each object j, the following weighted utilities, summed across the attributes i:

$$MAU_j = \sum_i w_i u_{ij} ,$$

where w_i is the relative importance of the i'th attribute and u_{ij} is the utility of the j'th object on the i'th attribute. For example, when choosing a car, w_i might be the importance of design, and u_{ij} would indicate how beautifully designed car j is. The theory prescribes that you choose the car with the largest MAUT. While this model is the most common, variants exist which incorporate additional features such as uncertainty, multiplicativity (rather than additivity) of the weighted utilities, time factors, and the possibility that your choice will affect others (293).

MAUT is a decision aid strongly grounded in theory. The axioms of the theory lead to the models, to methods for measuring the utilities and weights, and to specified tests which show which of the models is applicable. MAUT models have been extensively developed in the last five years (94, 95, 96, 141, 143, 233, 234). If these sources are too technical, try the review papers by MacCrimmon (186), Fischer (86, 88), von Winterfeldt & Fisher (296), Humphreys (131), and Huber (129a).

ASSESSMENT TECHNIQUES The first step in constructing a MAU is to list the attributes. Techniques for doing this are rarely discussed. Among those who have faced the problem, some have used the Delphi technique (e.g., 102, 211). Humphreys & Humphreys (132) suggested using George Kelly's repertory grid technique. Dalkey, Lewis & Snyder (65) proposed evaluating diverse problems (e.g., job choice, modes of transportation) not on the basis of their apparent attributes but on a common set of attributes reflecting quality of life (e.g., security, fun, freedom). Beach, et al (23) described an extensive interviewing technique, involving several interactions with different decision makers, to arrive at a list of attributes.

It seems obvious that the ommission of an important attribute can seriously alter the results of a MAUT application. However, Aschenbrenner & Kasubek (12) found reasonably similar results for preferences among apartments from MAU analyses based on two different, only partially overlapping sets of attributes.

Weights and utilities can be assessed either directly or indirectly. Direct approaches, which are simple but not theoretically justified, include ranking or rating scales, or just asking the assessor for the relevant numbers. For utilities, the assessor may be presented with graph paper and asked to sketch a curve. Utility functions may also be derived by constructing indifference curves for pairs of variables (189, 190); these methods are lengthy, tedious, and clearly impractical when there are many variables. After two indifference curves for the same pair of variables are assessed, a "staircase" method can be used by the analyst to uncover the utility curves for each of the variables, assuming that the variables are value independent (see 156, p. 57-61).

Indirect methods are justified within the theory, but are exceedingly complex. They rely on a comparison between a gamble and a sure thing, and thus introduce probabilities into an otherwise riskless situation. For example, to assess the weight of one attribute from a set of 14 attributes describing apartments (such as number of bedrooms, general cleanliness, etc.), the analyst says, "Apartment A has the best (most preferred) level of all 14 attributes. Apartment B has the worst level of all 14 attributes. Apartment C has the best level on one attribute and the worst level on each of the other 13. State a probability p such that you are indifferent between receiving C for sure versus receiving a gamble wherein you will obtain A with probability p and B with probability (1-p). What is the value of p that makes you indifferent?" The value of p that you name is the weight; such a question must be asked for each attribute.

The two indirect methods for assessing utilities are similar to the indirect method for assessing weights, except that "Apartment C" now has an <u>intermediate</u> level for one alternative, and the worst level for all others. In the variable-probability method, as with assessing weights, the task is to name a probability that makes the sure thing (Apartment C) indifferent to the gamble. In the fixed-probability method, the gamble's probabilities are held constant at (1/2, 1/2), and the assessor must name that intermediate value on one attribute of the sure thing which leads to indifference. In either case, one answer gives only one point on the utility curve, so that several responses are required to estimate its shape, for each attribute.

Kneppreth et al (156) have written an excellent review of the methods for assessing utilities, explaining each method in detail, noting advantages and disadvantages, and referencing relevant research. That research has been unsystematic and allows no clear conclusions. Perhaps future researchers should model their work on a study by Vertinsky & Wong (289). Comparing an indifference curve method with the indirect fixed-probability method, they looked at test-retest reliability and a host of other indices, including the acceptance of particular rationality axioms, realism of the task, confidence in the method, bias in the interpretation of probability, and a measure of the width of an indifference band across the variables. They found that the indirect method was more reliable and easier for the subjects, while the indifference curve technique predicted more subsequent choices.

ISSUES In MAUT, two issues are paramount. The first is: Is it valid? Early research in the use of MAUT frequently involved correlating the results of the model with unaided wholistic judgments of the same situations made by the same subjects (e.g., 130, 132, 294, and earlier papers referenced in the reviews mentioned above). A high correlation between the model and the wholistic judgments, the usual result, was taken as evidence that the model was valid. This conclusion seems faulty to us. If unaided wholistic preferences are good enough to constitute criteria for a decision aid like MAUT, who needs the decision aid? Furthermore, a decade or more of research has abundantly documented that humans are quite bad at making complex unaided decisions (248); it could thus be argued that high correlations with such flawed judgments would suggest a <u>lack</u> of validity. More sophisticated approaches have been taken by Fischer (87), who showed greater agreement among three different decomposition procedures than among three different wholistic procedures, and by Newman (208), who proposed applying Cronbach's (64) theory of generalizability to the problem of validating MAUT techniques.

But most practitioners and theorists approach the validity question as follows: the theory specifies the models, the assessment procedures, and the tests for choosing which model applies. Thus if you accept the axioms (yes, I do want my choices to be transitive; I should not be swayed by irrelevant alternatives, etc.) and pass the tests, then you can be assured that you are doing the right thing. There is no remaining validity question.

The second issue concerns error. Indirect elicitation techniques for both weights and utilities are, as previously noted, quite complex, but theoretically justifiable. The direct methods, in contrast, seem easier, but are theoretically unjustified. If one assumes that the decision maker has underlying weights, utilities, and preferences, which approach, direct or indirect, elicits these underlying values with least error? Von Winterfeldt (293) discussed but did not resolve this issue. Practitioners can (and often do) perform sensitivity analyses (how much can I change this parameter before the decision changes?). Such sensitivity analyses will identify where potential problems of measurement exist, but not solve them.

The tests which are used to determine which MAUT model is applicable are equally complex. The test for additivity uses the weights derived from the indirect method. If the weights across all the attributes sum to 1.0, an additive model may be used. Otherwise, a multiplicative model is used. No error theory is available to tell you whether a sum of, say, 1.4, is "close enough" to 1.0 to justify an additive model. An alternative, and seemingly easier, test is available for additivity (see 296, p. 70). Unfortunately, no alternatives are available for two other necessary tests. These tests are for two kinds of utility independence (called "preferential independence" and "utility independence" by Keeney (142), and "WCUI" and SCUI" by others [see 296]). The following question, with reference to the location of the Mexico City airport (142), is just the starting point for these tests: "How many people seriously injured or killed per year, call that number x, makes you indifferent between the option: [x injured or killed and 2500 persons subjected to high noise levels] and the option: [one person injured or killed and 1,500,000 subjected to high noise level]?" Several such questions must be asked for each attribute and for all pairs of attributes. The frequent avoidance of these tests may not reflect laziness, but a genuine suspicion that using an unjustified model may lead to fewer errors than choosing a model on the basis of confused responses to complex questions such as these. As von Winterfeldt (293) has noted, "even after you go through the process of model elimination and selection, you will still have to make up your mind about the possible trade-offs between assessment error and modeling error" (p. 65).

The flavor of the indirect assessment methods and the three tests mentioned above may be appreciated by reading 54 pages of dialogue between an analyst (Keeney) and an expert as they evaluate alternatives for the production of electrical energy (144).

RECENT RESEARCH The "new look" in MAUT research is to explore its uses. Can it be done? What problems are encountered? What can be learned from applying MAUT? Gardiner & Edwards (102) showed that in a highly controversial issue (coastal land development) two groups of experts (developers and conservationists) showed notably less disagreement about the evaluation of proposed apartment buildings in their MAUT evaluations than in their wholistic evaluations. O'Connor (211) reported the difficulties in getting many experts to agree on evaluations of water quality while trying to (a) minimize the amount of experts' time needed for the evaluation, (b) eliminate redundant or strongly interrelated attributes, (c) cope with possible non-compensatory factors (if the water is loaded with arsenic, nothing else matters). Guttentag & Sayeki (110) used a MAUT technique to illuminate the cultural differences in values and beliefs about peace issues between Japanese and Americans. In one of two reports of real applications (i.e., working with clients who paid for the advice), Keeney observed the changes in a MAUT system after two years of use (145). In the second report, he described the complexities of deciding where and when to build a new airport in Mexico City (142). Additional proposals for applications of MAUT, without relevant data, have been made for the development of social indicators (258), military system effectiveness (287) and solid waste management (150). Finally, computer programs to aid elicitation of MAUT have been written (146).

Decision Analysis

The most general approach for systematically evaluating alternative actions is decision analysis, an approach developed largely at the Harvard Business School (221, 235) and two private contract research firms, the Stanford Research Institute (125), and Decisions and Designs, Inc. (49). In facing a new problem, the analyst lists the decision alternatives, constructs a model of their interrelations, assesses the probabilities of relevant contingencies, finds out what the decision maker wants and, finally, assays the expected value or utility of each alternative. To do this, decision analysts use a bag of tricks drawn from crafts such as operations research, Bayesian statistics, SEU and MAUT, which allow the analysts to, "in principle, address any decision problem with unimpeachable rigor" (49, p. 64). A common tool is the decision tree which diagrams the uncertain consequences arising from a decision.

Among the problems that have been given full-dress decision analyses are whether to seed hurricanes in hopes of reducing their intensity (126), how to establish planetary quarantine requirements for trips to Mars and Jupiter (127), what value nuclear power generating plants have for Mexico (261), and how to design export controls on computer sales to the Soviet Bloc (71). Many environmental impact statements, cost-benefit analyses and risk assessments constitute variants on decision analytic-methodology (55, 91, 198, 216).

Although many of these analyses are already highly sophisticated, the basic methodology is still developing--often in response to specific problems. Work in the last five years has increased our ability to evaluate decision trees efficiently (288), assess the value of decision flexibility (194), and understand how models approximate the processes they are intended to describe (276).

Some awareness of psychological issues can be found in decision analysis. One example attempts to use the best psychological scaling techniques for eliciting probability judgments (260). Another emphasizes communicating effectively with decision makers; the analyst is encouraged to develop a role "not too dissimilar to that of a psychoanalyst" (49, p. 9). Brown (48) raised a cognitive problem that warrants further examination.

He noted that decision analyses often fail to model responses to future events. As a result, when those future events actually occur, they are responded to in totally unanticipated ways, because in the flesh they look different than they did at the time of the analysis.

Man/Machine Systems

For years, one of the most promising areas in decision aiding has been the development of computerized aids for helping decision makers cope with complex problems. Systems designed to elicit MAUT appraisals fall into this category, as do the approaches described below.

REGRESSION APPROACHES Research within the regression paradigm has shown that people have difficulty both applying the judgmental policies they wish to implement and describing the policies they actually are implementing. Hammond and colleagues have developed computer-graphics systems to combat both of these problems (113a, 117). Since these techniques can describe the policies of several participants in a given situation, they have been used to resolve interpersonal and intergroup conflicts (39) and to facilitate policy formation at the societal level (2, 116).

Another major decision-aiding technique is bootstrapping, which replaces judges with algebraic models of their own weighting policies. Recent research has continued to demonstrate that these models perform as well as or better than the judges themselves (14, 68, 119, 202, 237, 307). Additional work promises to further enhance bootstrapping's usefulness. Einhorn (81, 82) showed how expert judgment and statistical techniques can incorporate poorly defined and hard to measure variables into judges' models. Dawes & Corrigan ( ) demonstrated that in most situations the criterion being judged could be predicted well by models with unit weights (see also,). These unit weighting results suggest that in many decision [illegible]

Corrigan (70) demonstrated that in most situations the criterion being judged could be predicted well by models with unit weights (see also 83, 297). These unit-weighting results suggest that in many decision settings, all the judge needs to know is what variables to throw into the equation, which direction (+ or -) to weight them, and how to add. Actually, Benjamin Franklin had this insight about unit weighted linear models back in 1772 (186, p. 27).

PIP One of the earliest proposals for sharing the decision-making load between the machine and the decision maker was (79) the Probabilistic Information Processing System (PIP). In situations where judges must revise their probabilities upon receipt of new information, the PIP system accepts the judges' subjective assessments of prior probabilities, and of the probability of each datum conditional on each hypothesis, and then aggregates them according to Bayes' theorem in order to produce posterior probabilities of the hypotheses. A review in 1971 (254) revealed an abundance of research on PIP; since then, however, the flood has receded. A few recent studies have discussed what to do when the data are not conditionally independent of one another and have examined how well subjects handle such data (74, 129, 266). A couple of interesting medical applications have been proposed (108, 109).

DYNAMIC SYSTEMS Some of the most ambitious interactive man/machine systems have been developed to handle dynamic decision-making situations. The problems studied by researchers in this area are extremely varied and the systems developed to solve them tend to be highly specific. However, a pattern of conceptualizing the task, developing the mathematics and software to handle it, and then validating the system in one or a series of experiments is common. As an example, a team at Perceptronics, Inc. has developed a highly sophisticated system to assist naval officers tracking "the elements of a simulated fishing fleet [one trawler and one iceberg] as it moves about in an expanse of ocean" (a task that vaguely resembles a futuristic version of Battleships) (67, p. 3-1). The system tracks the decision maker's responses continuously and uses utilities inferred from them to recommend maximum expected utility decisions (98). From an experiment testing the system with 12 Naval Reserve NCO's during four 90-minute sessions, Davis et al (67) concluded that it worked in realistic decision-making situations, was accepted by experienced operators, and markedly improved performance.

Such systems may be designed either as products that will actually work in some field situation or as research tools. Perhaps because of their expense, most products have been designed to solve specific military problems with no civilian analog (although readers concerned about the possible presence of Soviet frogpersons in their bathtub or swimming pool might want to consult Irving [133]). It is difficult for the non-expert to judge the validity of these systems and the acceptability of their advice.

With systems designed for research purposes, a critical issue is the tradeoff between realism and generality. One strategy is to design systems whose complexity begins to approach that found in the real world--at the risk of investing too much of available resources in the machine and too little in understanding how people use it. Some human factors questions worth studying are (a) how do variations in the basic system (e.g., different instructions or information displays) affect people's performance? (b) how do person and machine errors interact? (c) how should machine output be adjusted to different decision makers' cognitive styles and work paces (170, 171)? and (d) when do people heed the machine's advice (111 , 112)?

Another problem with these systems is that their very complexity makes it difficult to compare results from one research context to the next. Perhaps the only way to do that is to interpret the results in terms of basic psychological (judgmental) phenomena. If that tack is taken, then one might ask whether the development of general behavioral principles would not be served best by using a number of simpler, cheaper and more flexible systems, such as the tactical and negotiations game used by the Streuferts and colleagues (e.g., 269). Research showing why man/machine systems should be adopted might provide a more convincing case than the demonstration in a complex simulation that decision makers do better with the machine's help. The skeptic may argue that such demonstrations merely show that one can design a simulated task in which it helps to have machine assistance.

Using Decision Aids

Do decision makers use these sophisticated techniques? Bootstrapping is now being applied for a variety of repeated decisions. On the other hand, apparently few, if any, PIP systems are operational today despite the mass of research refining its methodology. For most aids, a clear picture is hard to come by. In the scientific literature one can find demonstration projects showing a procedure's viability. However, when a technique passes the test of getting someone to pay for it, the result typically becomes proprietary. For reasons of national or industrial security, the details of such projects are not divulged, nor are the decision makers' responses to them. Most overviews by those in the decision aiding business understandably tend to be quite optimistic.

Brown (47, 49), however, has presented an insightful discussion of factors that may limit decision makers' receptiveness to decision analysis and presumably to other techniques as well. One is the fact that decision makers often employ an analyst to reduce the uncertainty in a problem situation, not to acknowledge and quantify it. Another source of resistance is the absence of top-level decision makers familiar with the technique; a third is the bad experiences of decision makers who try to solo on the technique without proper training. Brown, Kahr & Peterson (49) suggested that decision analysis is a clinical skill that should only be practiced after internship with an expert.

Another problem is that decision makers may, even after careful coaching, reject the basic conception (e.g., the axioms) on which the aids are based. Protocols of conversations between analysts and decision makers leave the impression that decision makers are under considerable pressure to adopt the analyst's perspective. It is debatable whether satisfaction with the results of such an analysis show that the analyst has really answered the decision maker's needs. Conrath (58) and Reeser (227) found that decision makers reject decision analysis (and related techniques) for being overly complicated and divorced from reality. Individuals who may accept the assumptions of such analysis may still reject their logical implications if they are unintuitive or too difficult to explain and justify to others.

A problem discussed earlier is whether decision makers can provide the required probability, utility and modeling judgments. Because of the vagaries of such judgments, the decision aider runs the risk of grinding through highly sophisticated analyses on inputs of very little value. Certainly "garbage in--garbage out" applies to decision aiding--with the particular danger that undue respect may be given to garbage produced by high-powered and expensive grinding. Relatively little is known about the sensitivity of decision aids to errors in elicitation and problem structuring. Von Winterfeldt & Edwards (294a) have proven that under very general conditions probability and utility estimates can be somewhat inaccurate without leading to appreciably suboptimal decisions. Their proof is applicable to the case where decision options are continuous (e.g., invest X dollars). However, Lichtenstein, Fischhoff & Phillips (175) have shown how a moderate error in probability estimation can lead to a substantial decrease in expected utility when the decision options are discrete (e.g., operate vs. don't operate). Von Winterfeldt & Edwards (295) have identified a large class of errors which can lead to large expected losses and are extremely difficult to detect. They arise from the selection of dominated decision alternatives as the result of inappropriately modeling the decision problem.

How much is a decision aid worth? This difficult question is typically answered with arguments why aids should, in principle, be worth the resources invested in them. Recently, Watson & Brown (303) provided enlightenment with a formal model for performing a decision analysis of a decision analysis. The model is accompanied by three case studies (304) that highlight the difficulties of performing a hindsightful analysis. Ironically, the greatest value of two of these analyses came from their contribution to organizational processes (reduction of controversy and improvement of communication), considerations that were left out of the formal model for the sake of simplicity.

CONCLUSION

One reason for the vitality of the research described here is the increased importance of deliberative decision making in our daily lives. In a non-traditional society individuals must rely on their analytical resources rather than habit in guiding their affairs. A rapidly changing and interrelated world cannot allow itself the luxury of trial and error as it attempts to cope with problems like nuclear power and natural hazard management. Economists, engineers, operations researchers, decision analysts and others are developing sophisticated procedures for these problems. It is our job as psychologists to remind them of the human component in implementing these techniques and explaining their conclusions to the public--in particular to point out the errors that may arise from judgmental biases. We must help the public to make its private decisions and to develop a critical perspective on those decisions made in its behalf.

Literature Cited[1]

1. Abelson, R. P. 1976. Script processing in attitude formation and decision making. In *Cognition and Social Behavior*, ed. J. S. Carroll, J. W. Payne. Hillsdale, N. J.: Lawrence Erlbaum Assoc. in press

2. Adelman, L., Stewart, T. R., Hammond, K. R. 1975. A case history of the application of social judgment theory to policy formulation. *Policy Sci*. 6:137-59

3. Alker, H. A., Hermann, M. G. 1971. Are Bayesian decisions artificially intelligent? The effect of task and personality on conservatism in processing information. *J. Pers. Soc. Psychol.* 19:31-41

4. Allais, P. M. 1953. The behavior of rational man in risk situations--A critique of the axioms and postulates of the American School. *Econometrika* 21:503-46

5. Anderson, N. H. 1972. Looking for configurality in clinical judgment. *Psychol. Bull*. 78:93-102

6. Anderson, N. H. 1974. Algebraic models in perception. In *Handbook of Perception*, ed. E. C. Carterette, M. P. Friedman, pp. 215-98 New York: Academic Press 556 pp.

7. Anderson, N. H. 1974. Information integration theory: A brief survey. In *Measurement, Psychophysics, and Neural Information Processing,* ed. D. H. Krantz, R. C. Atkinson, R. D. Luce, P. Suppes, 2:236-305. San Francisco: Freeman 468 pp.

8.

---

[1] To conserve space, frequently cited sources have been abbreviated as follows: JEP (*Journal of Experimental Psychology*); OBHP (*Organizational Behavior and Human Performance*).

8. Armelius, B., Armelius, K. 1974. Utilization of redundancy in multiple-cue judgments: Data from a suppressor variable task. Am. J. Psychol. 87:385-92

9. Armelius, K., Armelius, B. 1976. The effect of cue-criterion correlations, cue intercorrelations and the sign of the cue intercorrelation on performance in suppressor variable tasks. OBHP in press

10. Armstrong, J. S., Denniston, W. B. Jr., Gordon, M. M. 1975. The use of the decomposition principle in making judgments. OBHP 14:257-63

11. Armstrong, G. M., Kendall, C. L., Russ, F. A. 1975. Applications of consumer information processing research to public policy issues. Commun. Res. 2:232-45

12. Aschenbrenner, K. M., Kasubek, W. 1976. Convergence of multi-attribute evaluations when different sets of attributes are used. In Proceedings of the Fifth Research Conference on Subjective Probability, Utility, and Decision Making, ed. H. Jungermann, G. de Zeeuw, in press

13. Ashton, R. H. 1974. Cue utilization and expert judgments: A comparison of independent auditors with other judges. J. Appl. Psychol. 59:437-44

14. Ashton, R. H. 1975. User prediction models in accounting: An alternative use. Acctg. Rev. 50:710-22

15. Barclay, S., Beach, L. R., Braithwaite, W. P. 1971. Normative models in the study of cognition. OBHP 6:389-413

16. Bar-Hillel, M. 1973. On the subjective probability of compound events. OBHP 9:396-406

17. Bar-Hillel, M. 1974. Similarity and probability. OBHP 11:277-82

18. Barron, F. H. 1974. Behavioral decision theory: A topical bibliography for management scientists. Interfaces 5:56-62

19. Barron, F. H., Mackenzie, K. D. 1973. A constrained optimization model of risky decisions. J. Math. Psychol. 10:60-72

20. Bauer, M. 1971. Accuracy and congruence in estimations of probabilities and odds from binomial distributions. Umeå Psychol. Rep. 36. Umeå, Sweden: Univ. of Umeå

21. Bauer, M. 1973. Inference strategies in Bayesian tasks not requiring high scale-level responses. Umeå Psychol. Rep. 61. Umeå, Sweden: Univ. of Umeå

22. Beach, L. R. 1974. A note on the intrasubject similarity of subjective probabilities obtained by estimates and by bets. OBHP 11:250-252

23. Beach, L. R., Townes, B. D., Campbell, F. L., Keating, G. W. 1976. Developing and testing a decision aid for birth planning decisions. OBHP 15:99-116

24. Becker, G. M., McClintock, C. G. 1967. Value: Behavioral decision theory. Ann. Rev. Psychol. 18:239-86

25. Berkson, J., Magath, T. B., Hurn, M. 1940. The error of estimate of the blood cell count as made with the hemocytometer. Am. J. Physiology 128:309-23

26. Berl, J., Lewis, G., Morrison, R. S. 1976. Alternative models of choice in important and nonrepetitive situations. See Ref. 1

27. Bettman, J. R. 1971. A graph theory approach to comparing consumer information processing models. Mgmt. Sci. 18:114-28

28. Bettman, J. R. 1974. Toward a statistics for consumer decision net models. J. Consumer Res. 1:71-80

29. Bettman, J. R. 1975. Issues in designing consumer information environments. J. Consumer Res. 2:169-77

30. Bettman, J., Capon, N., Lutz, R. 1975. Multiattribute measurement models and multiattribute attitude theory: A test of construct validity. J. Consumer Res. 1:1-15

31. Bettman, J. R., Jacoby, J. 1975. Patterns of processing in consumer information acquisition. Papers in Consumer Psychol. No. 150. West Lafayette, Indiana: Purdue University

32. Birnbaum, M. H. 1976. Intuitive numerical prediction. Am. J. Psychol. in press

33. Björkman, M. 1973. Inference behavior in nonmetric ecologies. In Human Judgment and Social Interaction, ed. L. Rappoport, D. A. Summers, pp. 144-68. New York: Holt, Rinehart & Winston. 403 pp.

34. Bond, N. A. Jr. 1974. Basic strategy and expectation in casino blackjack. OBHP 12:413-28

35. Braithwaite, A. 1974. A note comparing three measures of subjective probability, their validity and reliability. Acta Psychol. 38:337-42

36. Brehmer, B. 1971. Subjects' ability to use functional rules. Psychonom. Sci. 24:259-60

37. Brehmer, B. 1973. Note on clinical judgment and the formal characteristics of clinical tasks. Umeå Psychol. Rep. 77. Umeå, Sweden: University of Umeå

38. Brehmer, B. 1974. Hypotheses about relations between scaled variables in the learning of probabilistic inference tasks. OBHP 11:1-27

39. Brehmer, B. 1976. Social judgment theory and the analysis of interpersonal conflict. Psychol. Bull. in press

40. Brehmer, B., Kuylenstierna, J., Liljergren, J. 1974. Effects of function form and cue validity on subjects' hypotheses in probabilistic inference tasks. OBHP 11:338-54

41. Brehmer, B., Kuylenstierna, J., Liljergren, J. 1975. Effects of information about the probabilistic nature of the task on learning of uncertain inference tasks. Umeå Psychol. Rep. 90. Umeå, Sweden: University of Umeå

42. Brehmer, B., Quarnstrom, G. 1976. Information integration and subjective weights in multiple-cue judgments. OBHP in press

43. Brewer, J. K., Owen, P. W. 1973. A note on the power of statistical tests in the Journal of Educational Measurement. J. Educ. Meas. 10:71-4

44. Brickman, P. 1972. Optional stopping on ascending and descending series. OBHP 7:53-62

45. Brickman, P., Pierce, S. M. 1972. Estimates of conditional probabilities of confirming versus disconfirming events as a function of inference situation and prior evidence. JEP 95:235-7

46. Brooks, W. N., Doob, A. N. 1975. Justice and the jury. J. Soc. Issues 31:171-82

47. Brown, R. V. 1971. Marketing applications of personalist decision analysis. MSI Field Res. Proj. Rep. P-55. Cambridge: Mgmt. Sci. Inst.

48. Brown, R. V. 1975. Modeling subsequent acts for decision analysis. DDI Tech. Rep. 75-1. McLean, Va.: Decisions & Designs, Inc.

49. Brown, R. V., Kahr, A. S., Peterson, C. 1974. Decision Analysis for the Manager. New York: Holt, Rinehart & Winston 618 pp.

50. Brown, T. A., Shuford, E. H. 1973. Quantifying uncertainty into numerical probabilities for the reporting of intelligence. RAND Rep. 1185-ARPA Santa Monica: Rand Corp.

51. Buckhout, R. 1974. Eyewitness testimony. Sci. Am. 231:23-31

52. Castellan, N. J. Jr. 1972. The analysis of multiple criteria in multiple-cue judgment tasks. OBHP 8:242-61

53. Castellan, N. J. Jr. 1973. Comments on the "lens model" equation and the analysis of multiple-cue judgment tasks. Psychometrika 38:87-100

54. Castellan, N. J. Jr. 1976. Decision making with multiple probabilistic cues. In Cognitive Theory, ed. N. J. Castellan, Jr., D. B. Pisoni, G. R. Potts. Vol. 2. Hillsdale, N. J.: Lawrence Erlbaum Assoc. in press

55. Coates, J. F. 1976. The role of formal models in technology assessment. Tech. Forecasting Soc. Change in press

56. Cohen, J. 1962. The statistical power of abnormal-social psychological research. J. Abnorm. Soc. Psychol. 65:145-53

57. Cohen, J., Chesnick, E. I., Haran, D. 1972. A confirmation of the inertial-ψ effect in sequential choice and decision. Br. J. Psychol. 63:41-6

58. Conrath, D. W. 1973. From statistical decision theory to practice: Some problems with the transition. Mgmt. Sci. 19:873-83

59. Cook, R. L. 1974. An interactive and iterative approach to computer-aided policy capturing. Prog. Res. Hum. Judg. Soc. Interact. Rep. 64. Boulder: Inst. Behav. Sci., Univ. Colorado

60. Coombs, C. H. 1975. Portfolio theory and the measurement of risk. In Human Judgment and Decision Processes, ed. M. F. Kaplan, S. Schwartz. pp. 64-83. New York: Academic Press 325 pp.

61. Coombs, C. H., Huang, L. C. 1974. Tests of the betweenness property of expected utility. MMPP Rep. 74-13. Ann Arbor: Univ. Michigan

62. Corbin, R., Marley, A. A. 1974. Random utility models with equality: An apparent but not actual generalization of random utility models. J. Math. Psychol. 11:274-93

63. Corbin, R. M., Olson, C. L., Abbondanza, M. 1975. Context effects in optional stopping decisions. OBHP 14:207-16

64. Cronbach, L. J., Gleser, G., Nanda, H., Rajaratnam, N. 1972. The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles. New York: Wiley

65. Dalkey, N. C., Lewis, R., Snyder, D. 1970. Measurement and analysis of the quality of life: With exploratory illustrations of applications to career and transportation choices. RAND RM-6228-DOT. Santa Monica: Rand Corp.

66. Davenport, W. G., Middleton, M. A. 1973. Expectation theories of decision making for duplex gambles. Acta Psychol. 37:155-72

67. Davis, K. B., Weisbrod, R. L., Freedy, A., Weltman, G. 1975. Adaptive computer aiding in dynamic decision processes: An experimental study of aiding effectiveness. Tech. Rep. PTR-1016-75-5. Woodland Hills, Calif.: Perceptronics, Inc.

68. Dawes, R. M. 1971. A case study of graduate admissions: Applications of three principles of human decision making. Am. Psychologist 26:180-88

69. Dawes, R. M. 1976. Shallow psychology. See Ref. 1

70. Dawes, R. M., Corrigan, B. 1974. Linear models in decision making. Psychol. Bull. 81:95-106

71. DDI 1973. Computer sale to the Soviet bloc. Tech. Rep. 73-4. McLean, Va.: Decisions & Designs, Inc.

72. Dershowitz, A. M. 1968. Psychiatry in the legal process: "Knife that cuts both ways" Judicature 51:370-77

73. Dillon, J. L. 1971. An expository review of Bernoullian decision theory. Rev. Mktg. Agri. Econ. 39:3-80

73a. Domas, P. A., Goodman, B. C., Peterson, C. R. 1972. Bayes's theorem: Response scales and feedback. Eng. Psychol. Lab. Tech. Rep. 037230-5-T. Ann Arbor: Univ. Michigan

74. Domas, P. A., Peterson, C. R. 1972. Probabilistic information processing systems: Evaluation with conditionally dependent data. OBHP 7:77-85

75. Donnell, M. L., DuCharme, W. M. 1975. The effect of Bayesian feedback on learning in an odds estimation task. OBHP 14:305-13

76. DuCharme, W. M., Donnell, M. L. 1973. Intrasubject comparison of four response modes for "subjective probability" assessment. OBHP 10:108-17

77. Ebert, R. J. 1972. Human control of a two-variable decision system. OBHP 7:237-64

78. Edwards, W. 1961. Behavioral decision theory. Ann. Rev. Psychol. 12:473-98

79. Edwards, W. 1962. Dynamic decision theory and probabilistic information processing. Hum. Factors 4:59-73

79a. Edwards, W. 1972. Application of research on cognition to man-machine system design. Eng. Psychol. Lab. Rep. 010342-1-F. Ann Arbor: Univ. Michigan

80. Edwards, W. 1975. Comment. J. Am. Statist. Assoc. 70:291-3

81. Einhorn, H. J. 1972. Expert measurement and mechanical combination. OBHP 7:86-106

82. Einhorn, H. J. 1974. Cue definition and residual judgment. OBHP 12:30-49

83. Einhorn, H. J., Hogarth, R. M. 1975. Unit weighting schemes for decision making. OBHP 13:171-92

84. Einhorn, H. J., Koelb, C. 1976. Psychometric study of literary critical judgment. Grad. Sch. Bus. Working Paper. Chicago: Univ. Chicago

85. Ellsberg, D. 1961. Risk, ambiguity, and the Savage axioms. Qtr. J. Econ. 75:643-9

86. Fischer, G. W. 1975. Experimental applications of multi-attribute utility models. In Utility, Probability, and Human Decision Making, ed. D. Wendt, C. A. J. Vlek, pp. 7-46. Dordrecht, The Netherlands: Reidel 418 pp.

87. Fischer, G. W. 1972. Four methods for assessing multi-attribute utilities: An experimental validation. Eng. Psychol. Lab. Tech. Rep. 037230-6-T. Ann Arbor: Univ. of Michigan

88. Fischer, G. W. 1976. Multidimensional utility models for risky and riskless choice. OBHP in press

89. Fischhoff, B. 1975. Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. JEP:Hum. Perc. Perf. 1:288-99

90. Fischhoff, B. 1976. Attribution theory and judgment under uncertainty. In New Directions in Attribution Research, ed. J. H. Harvey, W. J. Ickes, R. F. Kidd. Hillsdale, N. J.: Lawrence Erlbaum Assoc. in press

91. Fischhoff, B. 1976. Cost-benefit analysis and the art of motorcycle maintenance. ORI Res. Mono. 16(1). Eugene: Oregon Res. Inst.

92. Fischhoff, B. 1976. Perceived informativeness of factual information. ORI Res. Bull. 16(3). Eugene: Oregon Res. Inst.

93. Fischhoff, B., Beyth, R. 1975. "I knew it would happen"--remembered probabilities of once-future things. OBHP 13:1-16

94. Fishburn, P. C. 1970. Utility Theory for Decision Making, Pub. in Oper. Res. Series 18, ed. D. B. Hertz. New York: Wiley 234 pp.

95. Fishburn, P. C. 1974. von Neumann-Morgenstern utility functions on two attributes. Oper. Res. 22:35-45

96. Fishburn, P. C., Keeney, R. L. 1974. Seven independence concepts and continuous multiattribute utility functions. J. Math. Psychol. 11:294-327

97. Fontaine, G. 1975. Causal attribution in simulated versus real situations: When are people logical, when are they not? J. Pers. Soc. Psychol. 32:1021-9

98. Freedy, A., Weisbrod, R., Davis, K., May, D., Weltman, G. 1974. Adaptive computer aiding in dynamic decision processes: Adaptive decision models and dynamic utility estimation, Part I. Tech. Rep. PTR-1016-74-5(1). Woodland Hills, Calif.: Perceptronics, Inc.

99. Fryback, D. G., Goodman, B. C., Edwards, W. 1973. Choices among bets by Las Vegas gamblers: Absolute and contextual effects. JEP 98:271-8

100. Funaro, J. F. 1975. An empirical analysis of five descriptive models for cascaded inference. OBHP 14:186-206

101. Furby, L. 1973. Interpreting regression toward the mean in developmental research. Develop. Psychol. 8:172-9

102. Gardiner, P. C., Edwards, W. 1975. Public values: Multiattribute-utility measurement for social decision making. See Ref. 60, pp. 1-37

103. Gettys, C. F., Kelly, C. III, Peterson, C. R. 1973. The best guess hypothesis in multistage inference. OBHP 10:364-73

104. Gettys, C., Michel, C., Steiger, J. H., Kelly, C. W., Peterson, C. R. 1973. Multiple-stage probabilistic information processing. OBHP 10:374-87

105. Goodman, B. C. 1973. Direct estimation procedures for eliciting judgments about uncertain events. Eng. Psychol. Lab. Tech. Rep. 011313-5-T. Ann Arbor: Univ. Michigan

106. Graesser, C. C., Anderson, N. H. 1974. Cognitive algebra of the equation: Gift size = generosity x income. JEP 103:692-9

107. Green, P. E., Wind, Y. 1975. New way to measure consumers' judgments. Harvard Bus. Rev. 53:107-17

108. Greist, J. H., Gustafson, D. H., Stauss, F. F., Rowse, G. L., Laughren, T. P., Chiles, J. A. 1973. A computer interview for suicide-risk prediction. Am. J. Psychiatry 130:1327-32

109. Gustafson, D. H., Kestly, J. J., Greist, J. H., Jensen, N. M. 1971. Initial evaluation of a subjective Bayesian diagnostic system. Health Serv. Res. 6:204-13

110. Guttentag, M., Sayeki, Y. 1975. A decision-theoretic techniques for the illumination of cultural differences. J. Cross-Cultural Psychol. 6:203-17

111. Halpin, S. M., Johnson, E. M., Thornberry, J. A. 1973. Cognitive reliability in manned systems. IEEE Trans. Reliability R-22:165-70

112. Halpin, S. M., Thornberry, J. A., Streufert, S. 1973. The credibility of computer estimates in a simple decision making task. ONR Tech. Rep. 5. West Lafayette, Indiana: Purdue Univ.

113. Hammerton, M. 1973. A case of radical probability estimation. JEP 101:252-4

113a. Hammond, K. R. 1971. Computer graphics as an aid to learning. Science 172:903-8

114. Hammond, K. R. 1974. Human judgment and social policy. Prog. Res. Hum. Judg. Soc. Interact. Rep. 170. Boulder: Inst. Behav. Sci., Univ. Colorado

115. Hammond, K. R., Joyce, C. R. B., eds. 1975. Psychoactive Drugs and Social Judgment. New York: Wiley 278 pp.

116. Hammond, K. R., Stewart, T. R., Adelman, L., Wascoe, N. E. 1975. Report to the Denver city council and mayor regarding the choice of handgun ammunition for the Denver police department. Prog. Res. Hum. Judg. Soc. Interact. Rep. 179. Boulder: Inst. Behav. Sci., Univ. Colorado

117. Hammond, K. R., Stewart, T. R., Brehmer, B., Steinmann, D. O. 1975. Social judgment theory. See Ref. 60, pp. 271-312

118. Hammond, K. R., Summers, D. A. 1972. Cognitive control. Psychol. Rev. 79:58-67

119. Hamner, W. C., Carter, P. L. 1975. A comparison of alternative production management coefficient decision rules. Dec. Sci. 6:324-36

120. Hoffman, J., Peterson, C. R. 1972. A scoring rule to train probability assessors. Eng. Psychol. Lab. Tech. Rep. 037230-4-T. Ann Arbor: Univ. Michigan

121. Hogarth, R. M. 1974. Process tracing in clinical judgment. Behav. Sci. 19:298-313

121a. Hogarth, R. M. 1975. Cognitive processes and the assessment of subjective probability distributions. J. Am. Statist. Assoc. 70:271-94

122. Hogarth, R. M. 1975. Decision time as a function of task complexity. See Ref. 86, 321-38

123. Holzworth, R. J., Doherty, M. E. 1974. Inferences and predictions: normative vs. representative responding. Bull. Psychon. Soc. 3:300-02

124. Houle, A. 1973. Bibliography: Bayesian Statistics. Supplemented 1974-75. Ste-Foy, Quebec: Univ. of Laval

125. Howard, R. A., Matheson, J. E., Miller, K. E. 1976. Readings in Decision Analysis. Menlo Park: Stanford Res. Inst.

126. Howard, R. A., Matheson, J. E., North, D. W. 1972. The decision to seed hurricanes. Science 176:1191-1202

127. Howard, R. A., North, D. W., Pezier, J. P. 1975. A new methodology to integrate planetary quarantine requirements into mission planning, with application to a Jupiter orbiter. SRI Final Report NAS7-100. Stanford: Stanford Res. Inst.

128. Howell, W. C. 1972. Compounding uncertainty from internal sources. JEP 95:6-13

129. Howell, W. C., Gettys, G. F., Martin, D. W. 1971. On the allocation of inference functions in decision systems. OBHP 6:132-49

129a. Huber, G. P. 1974. Multi-attribute utility models: A review of field and field-like studies. Mgmt. Sci. 20:1393-1402

130. Huber, G. P., Daneshgar, R., Ford, D. L. 1971. An empirical comparison of five utility models for predicting job preferences. OBHP 6:267-82

131. Humphreys, P. 1976. Applications of multiattribute utility theory. See Ref. 12

132. Humphreys, P., Humphreys, A. 1975. An investigation of subjective preference orderings for multi-attributed alternatives. See Ref. 86, pp. 119-33

133. Irving, G. W. 1975. Alternative man/machine interface designs for swimmer defense systems. Integrated Sci. Corp. TM-75-36. Point Mugu, Calif.: Pacific Missile Test Cntr.

134. Jacoby, J. 1975. Perspectives on a consumer information processing research program. Commun. Res. 2:203-15

135. Jacoby, J. 1976. Consumer psychology: An octennium. Ann. Rev. Psychol. 27:331-58

136. Jensen, F. A., Peterson, C. R. 1973. Psychological effects of proper scoring rules. OBHP 9:307-17

137. Johnson, E. M., Cavanagh, R. C., Spooner, R. L., Samet, M. G. 1973. Utilization of reliability measurements in Bayesian inference: Models and human performance. IEEE Trans. Reliability R22:176-83

138. Kahneman, D., Tversky, A. 1972. Subjective probability: A judgment of representativeness. Cognitive Psychol. 3:430-54

139. Kahneman, D., Tversky, A. 1973. On the psychology of prediction. Psychol. Rev. 80:237-51

140. Kahneman, D., Tversky, A. 1975. Value Theory: An Analysis of Choices Under Risk. Presented at Conf. on Public Economics, Jerusalem, Israel

141. Keeney, R. L. 1971. Utility independence and preferences for multi-attributed consequences. Oper. Res. 19:875-93

142. Keeney, R. L. 1973. A decision analysis with multiple objectives: The Mexico City Airport. Bell J. Econ. Mgmt. Sci. 4:101-17

143. Keeney, R. L. 1974. Multiplicative utility functions. Oper. Res. 22:22-34

144. Keeney, R. L. 1975. Energy policy and value tradeoffs. IIASA Res. Memo. RM-75-76. Schloss Laxenburg, Austria: Int'l Inst. Appl. Sys. Analysis

145. Keeney, R. L. 1975. Examining corporate policy using multiattribute utility analysis. Sloan Mgmt. Rev. 17:63-76

146. Keeney, R. L., Sicherman, A. 1975. An interactive computer program for assessing and analyzing preferences concerning multiple objectives. IIASA Res. Memo. 75-12. Schloss Laxenburg, Austria: Int'l Inst. Appl. Sys. Analysis

147. Kelly, C. W., Peterson, C. R. 1971. Probability estimates and probabilistic procedures in current-intelligence analysis. IBM Rep. 71-5047. Gaithersburg, Md.: Int'l Bus. Mach. mach.

148. Kelly, C. W., Peterson, C. R. 1975. Decision theory research. DDI Tech. Rep. DT/TR 75-5. McLean, Va.: Decisions & Designs, Inc.

149. Kidd, J. B. 1970. The utilization of subjective probabilities in production planning. Acta Psychol. 34:338-47

150. Klee, A. J. 1971. The role of decision models in the evaluation of competing environmental health alternatives. Mgmt. Sci. 18B:52-67

151. Kleiter, G. D. 1975. Dynamic decision behavior: Comments on Rapoport's paper. See Ref. 86, pp. 371-80

152. Kleiter, G. D. 1975. Estimating the planning horizon in a multistage decision task. Psychol. Res. 38:37-64 1975

153. Kleiter, G., Gachowetz, H., Huber, D. 1976. Bibliography: Decision Making. Salzburg, Austria: Psychologisches Inst., Univ. Salzburg

154. Kleiter, G. D., Wimmer, H. 1974. Information seeking in a multistage betting game. Archiv für Psychol. 126:213-30

155. Knafl, K., Burkett, G. 1975. Professional socialization in a surgical specialty: Acquiring medical judgment. Soc. Sci. Med. 9:397-404

156. Kneppreth, N. P., Gustafson, D. H., Leifer, R. P., Johnson, E. M. 1974. Techniques for the assessment of worth. Tech. Paper 254. Arlington, Va.: Army Res. Inst.

157. Kozielecki, J. 1975. Psychologiczna Teoria Decyzji (Behavioral Decision Theory). Warszawa, PWN 352 pp. (Table of contents in Eng. & Russian)

158. Kozielecki. J. 1975. The internal representation of risky tasks. Polish Psychol. Bull. 6:115-21

159. Krantz, D. H., Atkinson, R. C., Luce, R. D., Suppes, P., eds. 1974. Contemporary Developments in Mathematical Psychology. Vol. 1 San Francisco: Freeman

160. Kunreuther, H. 1976. Limited knowledge and insurance protection. Public Policy in press

161. Kusyszyn, I. 1972. Psychology of gambling, risk-taking, and subjective probability: A bibliography. Journal Supplement Abstract Service. Catalog of selected documents in Psychol. 2:7

162. Kusyszyn, I. 1973. Gambling, risk-taking and personality: A bibliography. Int'l J. Addictions 8:173-90

163. Langer, E. J. 1975. The illusion of control. J. Pers. Soc. Psychol. 32:311-28

164. Langer, E. J., Roth, J. 1975. Heads I win, tails it's chance: The illusion of control as a function of the sequence of outcomes in a purely chance task. J. Pers. Soc. Psychol. 32:951-55

165. Lee, W. 1971. Decision Theory and Human Behavior. New York: Wiley 352 pp.

166. Leon, M., Anderson, N. H. 1974. A ratio rule from integration theory applied to inference judgments. JEP 102:27-36

167. Levin, I. P. 1974. Averaging processes and intuitive statistical judgments. OBHP 12:83-91

168. Levin, I. P. 1976. Information integration in numerical judgments and decision processes. JEP:Gen. 104:39-53

169. Levine, J. M., Samet, M. G. 1973. Information seeking with multiple sources of conflicting and unreliable information. Hum. Factors 15:407-19

170. Levine, J. M., Samet, M. G., Brahlek, R. E. 1975. Information seeking with limitations on available information and resources. Hum. Factors 17:502-13

171. Levit, R. A., Alden, D. G., Erickson, J. M., Heaton, B. J. 1974. Development and application of a decision aid for tactical control of battlefield operations. Vol. 2. ARI DAHC 19-73-C-0069. Minneapolis: Honeywell

172. Libby, R. 1975. The use of simulated decision makers in information evaluation. Acctng. Rev. 50:475-489

173. Lichtenstein, S., Earle, T., Slovic, P. 1975. Cue utilization in a numerical prediction task. JEP:Hum. Perc. Perf. 104:77-85

174. Lichtenstein, S. C., Fischhoff, B. 1976. Do those who know more also know more about how much they know? ORI Res. Bull. 16(1) Eugene: Oregon Res. Inst.

175. Lichtenstein, S. C., Fischhoff, B., Phillips, L. 1976. Calibration of probabilities: The state of the art. See Ref. 12

176. Lichtenstein, S. C., Slovic, P. 1971. Reversals of preference between bids and choices in gambling decision. JEP 89:46-55

177. Lichtenstein, S., Slovic, P. 1973. Response-induced reversals of preference in gambling: An extended replication in Las Vegas. JEP 101:16-20

178. Lindell, M. K., Stewart, T. R. 1974. The effects of redundancy in multiple-cue probability learning. Am. J. Psychol. 87:393-98

179. Lindman, H. R. 1971. Inconsistent preferences among gambles. JEP 89:390-7

180. Loftus, E. 1974. The incredible eyewitness. Psychol. Today 8:116-9

181. Lopes, L. 1976. Model based decision and judgment in stud poker. JEP:Gen. in press

182. Louviere, J. J. 1974. Predicting the evaluation of real stimulus objects from abstract evaluation of their attributes: The case of trout streams. J. Appl. Psychol. 59:572-7

183. Luce, R. D. 1959. Individual Choice Behavior. New York: Wiley

184. Lyon, D., Slovic, P. 1976. Dominance of accuracy information and neglect of base rates in probability estimation. Acta Psychol. in press

185. MacCrimmon, K. R. 1968. Descriptive and normative implications of the decision theory postulates. In Risk and Uncertainty, ed. K. Borch, J. Mossin. pp. 3-32. New York: St. Martin's Press. 455 pp.

186. MacCrimmon, K. R. 1973. An overview of multiple objective decision making. In Multiple Criteria Decision Making, ed. J. L. Cochrane, M. Zeleny. pp. 18-44. Columbia, S. C.: Univ. of So. Carolina Press. 816 pp.

187. MacCrimmon, K. R. 1974. Managerial decision making. In Contemporary Management: Issues and Viewpoints, ed. J. W. McGuire. pp. 445-95. Englewood Cliffs, N. J.: Prentice-Hall.

188. MacCrimmon, K. R., Larsson, S. 1976. Utility theory: Axioms versus "paradoxes". In Rational Decisions Under Uncertainty, special volume of Theory and Decision, ed. M. Allais, O. Hagen, in press

189. MacCrimmon, K. R., Siu, J. K. 1974. Making trade-offs. Dec. Sci. 5:680-704

190. MacCrimmon, K. R., Wehrung, D. A. 1975. Trade-off analysis: Indifference and preferred proportion. Fac. Commer. Bus. Admin. Working Paper 323. Vancouver, B. C.: Univ. of British Columbia

191. Marks, D. F., Clarkson, J. K. 1972. An explanation of conservatism in the bookbag-and-pokerchips situation. Acta Psychol. 36:145-60

192. Marks, D. F., Clarkson, J. K. 1973. Conservatism as non-Bayesian performance: A reply to DeSwart. Acta Psychol. 37:55-63

193. McCann, J. M., Miller, J. G., Moskowitz, H. 1975. Modeling and testing dynamic multivariate decision processes. OBHP 14:281-303

194. Merkhofer, M. W. 1975. Flexibility and decision analysis. Decision Anal. Prog. Res. Rep. EES-DA-75-1. Stanford: Stanford Univ.

195. Mertz, W. H., Doherty, M. E. 1974. The influence of task characteristics on strategies of cue combination. OBHP 12:196-216

196. Messick, D. M., Campos, F. T. 1972. Training and conservatism in subjective probability revision. JEP 94:335-7

197. Miller, P. M. 1971. Do labels mislead? A multiple cue study, within the framework of Brunswik's probabilistic functionalism. OBHP 6:480-500

198. Mishan, E. J. 1972. Cost-Benefit Analysis. New York: Praeger

198a. Monahan, J., Cummings, L. 1974. Prediction of dangerousness as a function of its perceived consequences. J. Crimin. Just. 2:239-42

199. Montgomery, H. 1976. A study of intransitive preferences using a think aloud procedure. See Ref. 12

200. Moskowitz, H. 1974. Effects of problem representation and feedback on rational behavior in Allais and Morlat-type problems. Dec. Sci. 5:225-42

201. Moskowitz, H. 1974. Regression models of behavior for managerial decision making. OMEGA, Int'l. J. Mgmt. Sci. 2:677-90

202. Moskowitz, H., Miller, J. G. 1972. Information and decision systems for production planning: An inter-disciplinary perspective. Inst. Res. Behav. Econ. Mgmt. Sci. paper 373. West Lafayette, Indiana: Purdue Univ.

203. Murphy, A. H. 1973. A new vector partition of the probability score. J. Appl. Meteor. 12:595-600

204. Murphy, A. H. 1974. A sample skill score for probability forecasts. Monthly Wthr. Rev. 102:48-55

205. Murphy, A. H., Winkler, R. L. 1970. Scoring rules in probability assessment and evaluation. Acta Psychol. 34:273-86

206. Murphy, A. H., Winkler, R. L. 1971. Forecasters and probability forecasts: Some current problems. Bull. Am. Meteor. Soc. 52:239-47

207. Murphy, A. H., Winkler, R. L. 1974. Probability forecasts: A survey of national weather service forecasters. Bull. Am. Meteor. Soc. 55:1449-53

208. Newman, J. R. 1975. Assessing the reliability and validity of multi-attribute utility procedures: An application of the theory of generalizability. SSRI Res. Rep. 75-7. Los Angeles: Univ. So. Calif.

209. Nickerson, R. S., Feehrer, C. E. 1975. Decision making and training: A review of theoretical and empirical studies of decision making and their implications for the training of decision makers. Tech. Rep. 73-C-0128-1. Orlando, Fla.: Naval Training Equip. Cntr. 210 pp.

210. Nisbett, R. E., Borgida, E. 1975. Attribution and the psychology of predictions. J. Pers. Soc. Psychol. 32:932-43

210a. Norman, K. L., Louviere, J. J. 1974. Integration of attributes in bus transportation: Two modeling approaches. J. Appl. Psychol. 59:753-8

211. O'Connor, M. F. 1973. The application of multiattribute scaling procedures to the development indices of water quality. Cntr. Math. Stud. Bus. Econ. Rep. 7339. Chicago: Univ. of Chicago

212. Ölander, F. 1975. Search behavior in non-simultaneous choice situations: Satisficing or maximizing. See Ref. 86, pp. 297-320

213. Payne, J. W. 1973. Alternative approaches to decision making under risk: Moments versus risk dimensions. Psychol. Bull. 80:439-53

214. Payne, J. W. 1976. Task complexity and contingent processing in decision making: An information search and protocol analysis. OBHP in press

215. Payne, J. W., Braunstein, M. L. 1971. Preferences among gambles with equal underlying distributions. JEP 87:13-18

216. Peskin, H. M., Seskin, E. P. 1973. Cost Benefit Analysis and Water Pollution Policy. Washington, D. C.: The Urban Inst. 325 pp.

217. Peterson, C. R., ed. 1973. Special Issue: Cascaded Inference. OBHP 10:315-432

218. Peterson, C. R., Beach, L. R. 1967. Man as an intuitive statistician. Psychol. Bull. 68:29-46

219. Pitz, G. F. 1974. Subjective probability distributions for imperfectly known quantities. In Knowledge and Cognition, ed. L. W. Gregg. pp. 29-41. New York: Wiley 321 pp.

220. Pollay, R. W. 1970. The structure of executive decisions and decision times. Admin. Sci. Qtrly. 15:459-71

221. Raiffa, H. 1968. Decision Analysis: Introductory Lectures on Choice Under Uncertainty. Reading, Mass.: Addison Wesley 309 pp.

222. Ramanaiah, N. V., Goldberg, L. R. 1976. Stylistic components of human judgment: The generality of individual differences. Appl. Psychol. Meas. in press

223. Rapoport, A. 1966. A study of human control in a stochastic multistage decision task. Behav. Sci. 11:18-32

224. Rapoport, A. 1975. Research paradigms for studying dynamic decision behavior. See Ref. 86, pp. 349-69

225. Rapoport, A., Burkheimer, G. J. 1971. Models for deferred decision making. J. Math. Psychol. 8:508-38

225a. Rapoport, A., Tversky, A. 1970. Choice behavior in an optimal stopping task. OBHP 5:105-20

226. Rapoport, A., Wallsten, T. S. 1972. Individual decision behavior. Ann. Rev. Psychol. 23:131-75

227. Reeser, C. 1971. The use of sophisticated analytical methods for decision making in the aerospace industry. MSU Bus. Topics 19:63-9

228. Restle, F. 1961. Psychology of Judgment and Choice. New York: Wiley

229. Ronen, J. 1973. Effects of some probability displays on choices. OBHP 9:1-15

230. Russo, J. E., Krieser, G., Miyashita, S. 1975. An effective display of unit price information. J. Marketing 39:11-19

231. Russo, J. E., Rosen, L. D. 1975. An eye fixation analysis of multialternative choice. Memory & Cognition 3:267-76

232. Savage, L. J. 1954. The Foundations of Statistics. New York: Wiley 294 pp.

233. Sayeki, Y. 1972. Allocation of importance: An axiom system. J. Math. Psychol. 9:55-65

234. Sayeki, Y., Vesper, K. H. 1973. Allocation of importance in a hierarchical goal structure. Mgmt. Sci. 19:667-75

235. Schlaifer, R. 1969. Analysis of Decisions Under Uncertainty. New York: McGraw-Hill. 729 pp.

236. Schmitt, N., Dudycha, A. 1975. A reevaluation of the effect of cue redundancy in multiple-cue probability learning. JEP 104:307-15

237. Schmidt, F. L., Marshall, R. L. 1973. Construction and use of a paramorphic representation of departmental policies in graduate admissions decision making. Journal Supplement Abstract Service, Catalog of selected documents in Psychol. 3:92

238. Schum, D. A. 1975. Contrast effects in inference: On the conditioning of current evidence by prior evidence. Res. Rep. Ser. 75-05. Houston, Texas: Rice Univ.

239. Schum, D. A. 1975. On the behavioral richness of cascaded inference models: Examples in jurisprudence. Res. Rep. Ser. 75-1. Houston, Texas: Rice Univ.

240. Seghers, R. C., Fryback, D. G., Goodman, B. C. 1973. Relative variance preferences in a choice-among-bets paradigm. Eng. Psychol. Lab. Tech. Rep. 011313-6-T. Ann Arbor: Univ. of Michigan

241. Selvidge, J. 1975. A three-step procedure for assigning probabilities to rare events. See Ref. 86, pp. 199-216

242. Shah, S. A. 1975. Dangerousness and civil commitment of the mentally ill: Some public policy consideration. Am. J. Psychiatry 132:501-5

243. Shanteau, J. 1972. Descriptive versus normative models of sequential inference judgment. JEP 93:63-8

244. Shanteau, J. 1975. An information-integration analysis of risky decision making. See Ref. 60, pp. 110-34

245. Sheridan, J. E., Richards, M. D., Slocum, J. W. 1975. Comparative analysis of expectancy and heuristic models of decision behavior. J. Appl. Psychol. 60:361-8

246. Shuford, E., Brown, T. A. 1975. Elicitation of personal probabilities and their assessment. Instruc. Sci. 4:137-88

247. Shulman, L. S., Elstein, A. S. 1975. Studies of problem solving, judgment, and decision making: Implications for educational research. In Review of Research in Education, ed. F. N. Kerlinger. 3:3-42 Itasca, Ill.: F. E. Peacock Pub. 305 pp.

248. Slovic, P. 1972. From Shakespeare to Simon: Speculations--and some evidence--about man's ability to process information. ORI Res. Mon. 12(2). Eugene: Oregon Research Institute

249. Slovic, P. 1972. Psychological study of human judgment: Implications for investment decision making. J. Finance 27:779-99

250. Slovic, P. 1975. Choice between equally-valued alternatives. JEP:Hum. Perc. Perf. 1:280-7

251. Slovic, P., Fischhoff, B., Lichtenstein, S. 1976. The certainty illusion. ORI Res. Bull. 16(4). Eugene: Oregon Research Institute

252. Slovic, P., Fischhoff, B., Lichtenstein, S. C. 1976. Cognitive processes and societal risk taking. See Ref. 15

253. Slovic, P., Kunreuther, H., White, G. F. 1974. Decision processes, rationality and adjustment to natural hazards. In Natural Hazards, Local, National and Global, ed. G. F. White. pp. 187-205. New York: Oxford Univ. Press 288 pp.

254. Slovic, P., Lichtenstein, S. 1971. Comparison of Bayesian and regression approaches to the study of information processing in judgment. OBHP 6:649-744

255. Slovic, P., MacPhillamy, D. J. 1974. Dimensional commensurability and cue utilization in comparative judgment. OBHP 11:172-94

256. Slovic, P., Tversky, A. 1974. Who accepts Savage's axiom? Behav. Sci. 19:368-73

257. Snapper, K. J., Fryback, D. G. 1971. Inferences based on unreliable reports. JEP 87:401-4

258. Snapper, K. J., O'Connor, M. F., Einhorn, H. J. 1974. Social indicators: A new method for indexing quality. Soc. Res. Gr. Tech. Rep. 74-4. Washington, D. C.: George Washington Univ.

259. Snapper, K. J., Peterson, C. R. 1971. Information seeking and data diagnosticity. JEP 87:429-33

260. Spetzler, C. S., Staël von Holstein, C.-A. S. 1975. Probability encoding in decision analysis. Mgmt. Sci. 22:340-58

261. SRI 1968. Decision analysis of nuclear plants in electrical system expansion. SRI Proj. 6496 Final Rep. Stanford: Stanford Res. Inst.

262. Stachowski, R. 1974. Effect of predecisional information integration strategy on cognitive conservatism. Polish Psychol. Bull. 5:17-23

263. Staël von Holstein, C.-A. S. 1971. An experiment in probabilistic weather forecasting. J. Appl. Meteor. 10:635-45

264. Staël von Holstein, C.-A. S. 1971. Two techniques for assessment of subjective probability distributions--an experimental study. Acta Psychol. 35:478-94

265. Staël von Holstein, C.-A. S. 1972. Probabilistic forecasting: An experiment related to the stock market. OBHP 8:139-58

266. Steiger, J. H., Gettys, C. F. 1972. Best-guess errors in multistage inferences. JEP 92:1-7

267. Stenson, H. H. 1974. The lens model with unknown cue structure. Psychol. Rev. 81:257-64

268. Stewart, T. R., Carter, J. E. 1973. POLICY: An interactive computer program for externalizing, executing, and refining judgmental policy. Prog. Res. Hum. Judg. Soc. Interact. Rep. 159. Boulder: Inst. Behav. Sci., Univ. Colorado

269. Streufert, S. C. 1973. Effects of information relevance on decision making in complex environments. Memory & Cognition 1:224-8

270. Sue, S., Smith, R. E., Caldwell, C. 1973. Effects of inadmissible evidence on the decisions of simulated jurors: A moral dilemma. J. Appl. Soc. Psychol. 3:345-53

271. Svenson, O. 1973. Analysis of strategies in subjective probability inferences as evidenced in continuous verbal reports and numerical responses. Psychol. Labs. Rep. 396. Stockholm, Sweden: Univ. Stockholm

272. Svenson, O. 1974. A note on think aloud protocols obtained during the choice of a home. Psychol. Labs. Rep. 421. Stockholm, Sweden: Univ. Stockholm

273. Svenson, O. 1975. A unifying interpretation of different models for the integration of information when evaluating gambles. Scand. J. Psychol. 16:187-92

274. Svenson, O., Montgomery, H. 1974. A frame of reference for the study of decision processes. Psychol. Labs. Rep. 409. Stockholm, Sweden: Univ. Stockholm

275. Swinth, R. L., Gaumnitz, J. E., Rodriguez, C. 1975. Decision making processes: Using discrimination nets for security selection. Dec. Sci. 6:439-48

276. Tani, S. N. 1975. Modeling and decision analysis. Dec. Anal. Prog. Res. Rep. EES-DA-75-3. Stanford: Stanford Univ.

277. Taylor, R. L., Wilsted, W. D. 1974. Capturing judgment policies: A field study of performance appraisal. Acad. Mgmt. J. 17:440-9

278. Teigen, K. H. 1974. Overestimation of subjective probabilities. Scand. J. Psychol. 15:56-62

279. Teigen, K. H. 1974. Subjective sampling distributions and the additivity of estimates. Scand. J. Psychol. 15:50-5

280. Tversky, A. 1972. Choice by elimination. J. Math. Psychol. 9:341-67

281. Tversky, A. 1972. Elimination by aspects: A theory of choice. Psychol. Rev. 79:281-99

282. Tversky, A. 1975. Assessing uncertainty. J. Roy. Statist. Soc. 36B:148-59

283. Tversky, A. 1975. On the Elicitation of Preferences: Descriptive and Prescriptive Considerations. Presented at Workshop on Dec. Making with Multiple Conflicting Objectives, IIASA, Schloss Laxenburg, Austria

284. Tversky, A., Kahneman, D. 1971. The belief in the "law of small numbers." Psychol. Bull. 76:105-10

285. Tversky, A., Kahneman, D. 1973. Availability: A heuristic for judging frequency and probability. Cognitive Psychol. 5:207-32

286. Tversky, A., Kahneman, D. 1974. Judgment under uncertainty: Heuristics and biases. Science 185:1124-31

287. Turban, E., Metersky, M. L. 1971. Utility theory applied to multivariate system effectiveness evaluation. Mgmt. Sci. 17B:817-28

288. Ulvila, J. W. 1975. A pilot survey of computer programs for decision analysis. DDI Tech. Rep. 75-2. McLean, Va.: Decisions & Designs, Inc.

289. Vertinsky, I., Wong, E. 1975. Eliciting preferences and the construction of indifference maps: A comparative empirical evaluation of two measurement methodologies. Socio-Econ. Plan. Sci. 9:15-24

290. Vlek, C. A. J. 1973. Coherence of human judgment in a limited probabilistic environment. OBHP 9:460-81

291. Vlek, C. A. J. 1973. The fair betting game as an admissible procedure for assessment of subjective probabilities. Br. J. Math. Statist. Psychol. 26:18-30

292. Vlek, C. A. J., Wagenaar, W. A. 1975. Judgment and Decision Under Uncertainty. Leiden, The Netherlands: Univ. Leiden 82 pp.

293. von Winterfeldt, D. 1975. An overview, integration, and evaluation of utility theory for decision analysis. SSRI Res. Rep. 75-9 Los Angeles: Univ. So. Calif.

294. von Winterfeldt, D., Edwards, W. 1973. Evaluation of complex stimuli using multi-attribute utility procedures. Eng. Psychol. Lab. Tech. Rep. 011313-2-T. Ann Arbor: Univ. Michigan

294a. von Winterfeldt, D., Edwards, W. 1973. Flat maxima in linear optimization models. Eng. Psychol. Lab. Tech. Rep. 011313-4-T. Ann Arbor: Univ. Michigan

295. von Winterfeldt, D., Edwards, W. 1975. Error in decision analysis: How to create the possibility of large losses by using dominated strategies. SSRI Res. Rep. 75-4. Los Angeles: Univ. So. Calif.

296. von Winterfeldt, D., Fischer, G. W. 1975. Multi-attribute utility theory: Models and assessment procedures. See Ref. 86, pp. 47-86

297. Wainer, H. 1976. Estimating coefficients in linear models: It don't make no nevermind. Psychol. Bull. 83:213-17

298. Wainer, H., Zill, N., Gruvaeus, G. 1973. Senatorial decision making: II. Prediction. Behav. Sci. 18:20-6

299. Wallsten, T. S. 1971. Subjectively expected utility theory and subjects' probability estimates: Use of measurement-free techniques. JEP 88:31-40

300. Wallsten, T. S. 1972. Conjoint-measurement framework for the study of probabilistic information processing. Psychol. Rev. 79:245-60

301. Wallsten, T. 1975. Using a conjoint measurement model to develop theory about probabilistic information processing. Psychometric Lab. Rep 127 (revised). Chapel Hill, N. C.: Univ. N. Carolina

302. Ward, W. M. 1975. Heuristic use or information integration in the estimation of subjective likelihood? Bull. Psychon. Soc. 6:43-6

303. Watson, S. R., Brown, R. V. 1975. Issues in the value of decision analysis. DDI Tech. Rep 75-9 McLean, Va.: Decisions & Designs, Inc.

304. Watson, S. R., Brown, R. V. 1975. Case studies in the value of decision analysis. DDI Tech. Rep. 75-10 McLean, Va.: Decisions & Designs, Inc.

305. Wheeler, G. E., Edwards, W. 1975. Misaggregation explains conservative inference about normally distributed populations. SSRI Res. Rep. 75-11. Los Angeles: Univ. So. Calif.

306. Wiggins, N. 1973. Individual differences in human judgments: A multivariate approach. See Ref. 33, pp. 110-42

307. Wiggins, N., Kohen, E. S. 1971. Man versus model of man revisited: The forecasting of graduate school success. J. Pers. Soc. Psychol. 19:100-6

308. Winkler, R. L. 1971. Probabilistic prediction: Some experimental results. J. Am. Statist. Assn. 66:675-85

309. Winkler, R. L., Murphy, A. H. 1973. Experiments in the laboratory and the real world. OBHP 10:252-70

310. Wise, J. A., Mockovak, W. P. 1973. Descriptive modeling of subjective probabilities. OBHP 9:292-306

311. Wohlstetter, A. 1974. Legends of the strategic arms race, Part I: The driving machine. Strategic Rev. pp. 67-92

311a. Wohlstetter, R. 1962. Pearl Harbor: Warning and Decision. Stanford: Stanford Univ. Press 422 pp.

312. Wright, P. L. 1973. Use of consumer judgment models in promotion planning. J. Marketing 37:27-33

313. Wright, P. 1974. The harassed decision maker: Time pressures, distractions and the use of evidence. J. Appl. Psychol. 59:555-61

314. Wright, P. 1974. The use of phased, noncompensatory strategies in decisions between multiattribute products. Grad. Sch. Bus. Res. Paper Ser. 223. Stanford: Stanford Univ.

315. Wright, W. F. 1975. Cognitive information processing models: An empirical study. Grad. Sch. Bus. Res. Paper Ser. 246 Stanford: Stanford Univ.

316. Wyer, R. S. 1974. Cognitive Organization and Change: An Information Processing Approach. Potomac, Md.: Lawrence Erlbaum 502 pp.

317. Wyer, R. S. 1976. An investigation of the relations among probability estimates. OBHP 15:1-18

318. Zagorski, M. A. 1975. Risky decision: attention effects or masking effects? Acta Psychol. 39:487-94

319. Zieve, L. 1966. Misinterpretation and abuse of laboratory tests by clinicians. Annals N. Y. Acad. Sci. 134:563-72