JOURNAL OF MATHEMATICAL PSYCHOLOGY 33, 397–427 (1989)

An Empirical Evaluation of Descriptive Models of Ambiguity Reactions in Choice Situations

SHAWN P. CURLEY

University of Minnesota

AND

J. FRANK YATES

University of Michigan

Ambiguity is uncertainty about an option's outcome-generating process, and is characterized as uncertainty about an option's outcome probabilities. Subjects, in choice tasks, typically have avoided ambiguous options. Descriptive models are identified and tested in two studies which had subjects rank monetary lotteries according to preference. In Study 1, lotteries involved receiving a positive amount or nothing, where P denotes the probability of receiving the nonzero amount. Subjects were willing to forego expected winnings to avoid ambiguity near $P = .50$ and $P = .75$. Near $P = .25$, a significant percentage of subjects exhibited ambiguity seeking, with subjects, on average, willing to forego expected winnings to have the more ambiguous option. The observed behavior contradicts the viability of a proposed lexicographic model. Study 2 tested four polynomial models using diagnostic properties in the context of conjoint measurement theory. The results supported a sign dependence of ambiguity with respect to the probability level P, such that subjects' preference orderings over ambiguity reversed with changes in P. This behavior was inconsistent with all the three-factor polynomial models investigated. Further analyses failed to support a variant of portfolio theory, as well. The implications of these results for the descriptive modeling of choice under ambiguity are discussed.

Suppose you feel stiffness and pain in your legs after walking several blocks. At the clinic, you are informed of two available treatments, Treatment A and Treatment B. You describe your choice, as to which treatment to accept, if either, by the tree structure shown in Fig. 1. The structure captures your beliefs that the possible outcomes at least partly depend upon your choice, and that the outcomes are uncertain. These features are components of most decision models under uncertainty that have been proposed. Your next step could be to use probabilities, indicated on the tree in Fig. 1, as subjective measures of your uncertainty about the chances that each treatment will be successful. A decision analysis would proceed in this fashion (Weinstein *et al.*, 1980; Winkler, 1972).

In structuring a choice in this fashion, we essentially treat the process whereby the outcomes are generated as analogous to a lottery, or chance, model with known

Reprint requests should be addressed to Dr. Shawn P. Curley, Department of Information and Decision Sciences, University of Minnesota, 271 19th Ave. South, Minneapolis, MN 55455.

0022-2496/89 $3.00

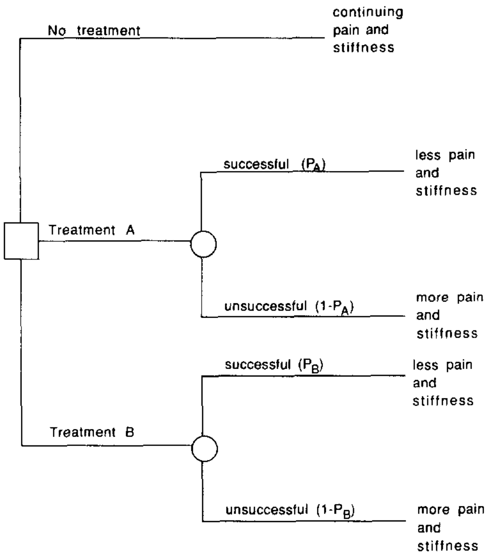

FIG. 1. Decision tree representation of a hypothetical clinical scenario with two available treatments, A and B (adapted from Curley, Eraker, and Yates, 1984).

canonical probabilities (Einhorn & Hogarth, 1986; Shafer & Tversky, 1985). We specify our uncertainty using subjective probabilities derived from a metaphor of a known, probabilistic outcome-generating process. In contrast, *ambiguity* is uncertainty about what outcome-generating process is appropriate or about how an outcome-generating process is operating. In that an uncertain process or competing processes typically imply differing outcome probabilities, ambiguity has been operationalized as an imprecision in the judged probabilities. Furthermore, whereas one's uncertainty about decision outcomes can be captured by probabilities, the uncertainty of ambiguity is inconsistent with even a qualitative probability measure (Ellsberg, 1961). Probabilities only partially measure the uncertainty of interest to decision makers.

Ambiguity, as a third feature of decision situations, underlies Keynes's (1921) argument in the following passage:

> But it seems that there may be another respect in which some kind of quantitative comparison between arguments is possible. This comparison turns upon a balance, not between the favourable and the unfavourable evidence, but between the *absolute* amounts of relevant knowledge and of relevant ignorance respectively. (p. 71)

Keynes's claim was that decisions differ in the amount of knowledge one has about the chances of the outcomes involved. Lack of evidence is one component which Ellsberg (1961) identified as denoting ambiguity. Ambiguity might also be introduced by the presence of conflicting evidence, unreliable evidence, or some other source of inherent uncertainty about how the outcomes will be generated. Thus, there may be more or better evidence regarding the outcome-generating process, and, consequently, about the outcome likelihoods for Treatment A than for Treatment B. This factor, slighted by the representation in Fig. 1, may be of concern to you, beyond the uncertainty measured by the probabilities.

The usual neglect of ambiguity in structuring decisions has significance due both to the prevalence of ambiguity in choice tasks and to the reactions of persons to ambiguity. That ambiguity is prevalent has not been disputed; rarely in a decision situation can we precisely specify probabilities for all the possible outcomes. A stated probability represents a compromise among possibilities. Equally verifiable is the existence of ambiguity reactions. Typically the reaction is in the form of *ambiguity avoidance*, which is a preference for the option having the least ambiguity, when given a choice between options differing only in their degrees of ambiguity. In certain circumstances, systematic *ambiguity seeking* might also obtain (Einhorn & Hogarth, 1986). Since Ellsberg (1961, 1963) first defined ambiguity reactions, as alluded earlier by Keynes (1921) and Knight (1921), numerous empirical studies have verified their conjecture in a variety of decision contexts (cf. Curley & Yates, 1985).

Of present interest are formal models describing observed reactions to ambiguity. First, consider the model of statistical decision theory, which, in its general form, can be summarized with the form $P * U$, involving two functions: a probability or weight distribution P, whether these probabilities be subjective or otherwise, and a utility or value function U (cf. Schoemaker, 1982 for instances of this general form). In that some subjects' behavior under ambiguity in certain situations can be described by the model of statistical decision theory, the model has offered a promising starting point for the development of descriptive models. The basic process implied by the model is that people make decisions as a tradeoff between likelihoods and values.

Still, for choices under ambiguity, the model, which is derived for situations having no ambiguity, is insufficient as a general description of ambiguity reactions (Ellsberg, 1961). If identifiable, a descriptive model could illuminate the processes whereby decision makers evaluate alternatives in, as well as provide insight into the structure of, realistic choice situations involving ambiguity. Section 1 outlines generalized forms of models which have been proposed to describe ambiguity reactions, focusing upon the underlying processes which they might represent. Section 2 presents two studies testing these models. Section 3 concludes with a general discussion of the status of descriptive choice models under ambiguity.

Proposed Descriptive Models

Five plausible three-factor models that generalize the statistical decision model (Model 1) are displayed in Table 1 as Models 2–6. These models all include an ambiguity factor as an addition to the factors comprising Model 1. Models 3–6 are all possible combination rules involving utility, probability, and ambiguity factors

TABLE 1

General Choice Models 1–6

Model 1.	Statistical Decision Theory	$P * U$
Model 2.	Lexicographic	$P * U, A$
Model 3.	Distributive	$(P + A) * U$
Model 4.	Distributive	$P * (U + A)$
Model 5.	Dual-Distributive	$P * U + A$
Model 6.	Multiplicative	$P * U * A$

Note: P = a probability or subjective weight function; U = a utility or value function; A = an ambiguity function.

and the two algebraic operations of multiplication (*) and addition (+) between factors which have the model $P * U$ as a special case. Model 2 incorporates the ambiguity factor within a lexicographic model.

As it turns out, most of the models in the literature are special cases of the models in Table 1 and were created in a similar manner, namely, by generalizing from the model of statistical decision theory. Also, with rare exception, the proposed models have not been accompanied by empirical verification. In this section, the models are first described; then consideration is given to differentiating among them. Citations for models which are special cases of each of the generalized forms in Table 1 are provided below (cf. Curley, 1986, for more detail on the relationships between the specific models and the generalized forms).

It is possible that none of the three-factor models is adequately descriptive. Only in the event that none of these models proves adequate would we wish to consider more complex generalizations with more than three factors, or other forms. Examples of more complex generalizations are also acknowledged below.

Lexicographic: $P * U, A$ (Roberts, 1963)

This model of ambiguity avoidance is a lexicographic decision rule: The less ambiguous option is only selected for options which evaluate to the same expectation as described by the statistical decision model. Ambiguity is the second dimension considered. The lexicographic process can be thought of as an orderly tie-breaking rule. When a tie is encountered, after considering likelihoods and values, the rule provides a criterion for making a selection. The presence of such a rule allows for easy, quick evaluation of alternatives which are equivalent in subjective expectation.

Distributive: $(P + A) * U$ (Ellsberg, 1961; Gardenfors, 1979; Gardenfors & Sahlin, 1982a, 1982b, 1983)

This distributive model of ambiguity reactions might describe a process in which the decision maker arrives at a composite measure of uncertainty, combining outcome and process uncertainty, before incorporating the value information. The process, for example, could result from a modification of the best-guess probability to account for a worse, or worst, possible case. Such a process would induce ambiguity avoidance as has been observed. In other situations or for other decision makers, a better, or best, possible case could be more compelling, leading to ambiguity-seeking behavior.

Distributive: $P * (U + A)$ (Smith, 1969)

The second distributive model is similar to the first as potentially arising from an additive modification process. In this case, ambiguity is perceived as affecting decisions through value modification with a utility, or disutility, being attached to the presence of ambiguity. The composite utility has value $(U + A)$. The value modification might be prompted by consideration of a utility for gambling, similar to that which has been applied for decisions under risk (Fishburn, 1980), or by a value–uncertainty interaction effect similar to that observed between value and judged probability (Lee, 1971; Slovic, 1966).

Dual-Distributive: $P * U + A$ (Toda & Shuford, 1965)

The dual-distributive model could also arise from the presence of a utility for gambling or process uncertainty. This model more closely resembles Fishburn's (1980) model. Whereas the distributive model $P * (U + A)$ applies the value of ambiguity to each outcome separately, the dual-distributive model describes a process whereby the ambiguity of the option, as a whole, is used in modifying the evaluation of the option. This is perhaps a more natural mechanism through which a utility for ambiguity might operate. Other processes whereby ambiguity is evaluated as a global aspect of the option and then applied to modify the option's evaluation may be similarly described with the dual-distributive form. For example, the measure could indicate that the decision maker is reserving part of his or her belief to an unattached state, in reaction to uncertainty about the outcomes' relative likelihoods (Fellner, 1961). Similar interpretations are given by Shafer (1976) to a "degree of belief" measure and by Zadeh (1978) to a "possibility" measure.

Multiplicative: $P * U * A$ (Fellner, 1961)

The multiplicative model may arise from a process similar to any of those described for the distributive and dual-distributive models. Thus, the decision maker might employ multiplicative composite uncertainty measures $(P * A)$, which are then combined with values; the decision maker might evaluate multiplicative composite utilities $(U * A)$, which are then weighted by the probabilities; or the decision maker might modify the expectation $(P * U)$, incorporating ambiguity more globally. The underlying processes may be equivalent to those described for the other models; only the form of the modifications differs.

More than Three Factors (Einhorn & Hogarth, 1985, 1988)

An example of a model which is more complex than those in Table 1 is presented by Einhorn and Hogarth (1985). The model was proposed primarily as a model of likelihood judgments for options involving ambiguity. One possible adaptation to the choice task can be summarized as $(P + A * P') * U$, where P' is an ambiguity-dependent probability function, in contrast to P, which is independent of the presence of ambiguity. For example, in Einhorn and Hogarth's model, $P' = (1 - P - P^{\beta})$, where β is a measure of the individual's attitude toward ambiguity in situations of the considered type and represents a differential weighting of probability values above and below P. This generalized form of their model is just one of many possible polynomial forms utilizing more than three factors. An alternative choice model, involving more than four factors, is described by Einhorn and Hogarth (1988).

Other Forms

Other models have been developed to describe decisions under risk, which can be generalized to choice under ambiguity in the same way that we have generalized the statistical decision theory model. One example is portfolio theory as described by Coombs and Meyer (1969) and Coombs and Huang (1970). Portfolio theory describes choice under risk via a preference plane over the two dimensions of expected value and risk. The variable of risk is left undefined. However, the model does propose several assumptions about the effects of changes in the option parameters upon perceived risk. Portfolio theory can be adapted to represent choice under ambiguity by including ambiguity as another parameter which influences the undefined variable of perceived risk. In other words, perhaps ambiguity is just another form, or aspect, of risk which can be incorporated into the theory.

Differentiating the Models

Which, if any, of the models in Table 1 best describes subjects' reactions to ambiguity in choice situations? The models which have been proposed in the literature, and which have been generalized to the forms in Table 1, differ in two respects. Each involves a different definition of the ambiguity function A used to modify Model 1; and the models differ in the composition rule for combining the function A with the functions P and U. These two differences are separable, representing two distinct descriptive issues: How is ambiguity operationally defined? and What is the model of choice? The present state of knowledge regarding each of these questions is incomplete and worthy of attention. Curley, Young, and Yates (1989) and Budescu and Wallsten (1985) describe experimental approaches for the first of these issues; the present studies address the second.

There is a benefit to separating the issues of defining the function A and of modeling choice behavior. By differentiation of the models without explicit specification of the ambiguity function A, the results of the investigation may be generalized beyond the specific operationalizations of ambiguity that have been used in the literature. This allows analyses among the more general versions of the models shown in Table 1, rather than specific versions of those models which incorporate particular characterizations of ambiguity, including those explicit in the literature.

Testing among models which contain an ambiguity factor A that is not defined appears problematic, but it is not untenable. The situation parallels several which have appeared in the literature. One example is that of models of choice which incorporate "risk" as a construct, in the absence of a well-characterized concept of risk (e.g., Coombs, Donnell, & Kirk, 1978; Coombs & Huang, 1970; Libby & Fishburn, 1977; Slovic, 1964). The approach is to extract some property, or properties, of the undefined concept about which there is consensus. The consensual properties are then included as assumptions in the testing of the model. In fact, authors have used a similar approach in handling the issue of defining risk. Properties of risk are used in developing possible models, which can then be axiomatized and examined in detail (e.g., Luce, 1980; Pollatsek & Tversky, 1970).

A similar strategy has been employed in demonstrations of the influence of representativeness upon judgments (Kahneman & Tversky, 1972, 1973; Tversky & Kahneman, 1971). Options are devised about which there is consensus as to their ordering with respect to representativeness. It is then demonstrated, using these options, how subjects' behavior depends upon the concept of representativeness, a concept which is not well-defined.

The present research involves two studies designed to illuminate the issue of descriptive models of choice under ambiguity. Study 1 provides a test of Model 2. The study was also designed to obtain approximate measures of the strength of ambiguity reactions using a choice paradigm. To date, such measures have only been elicited using pricing procedures (Becker & Brownson, 1964; Yates & Zukowski, 1976). However, a subject's behavior in a pricing task does not necessarily generalize to choice situations (Goldstein, 1985; Goldstein & Einhorn, 1987; Grether & Plott, 1979; Lichtenstein & Slovic, 1971). Study 2 examines the three-factor Models 3–6, using procedures derived from the theory of conjoint measurement. The study also addresses the described adaptation of portfolio theory for ambiguous choice situations.

EXPERIMENT

Both Studies 1 and 2 were conducted in a single experimental session with the same subject population. They involved similar stimuli and experimental tasks. The general method used in the experiment opens this section.

General Method

Subjects

Thirty-one undergraduates at the University of Michigan participated in fulfillment of a requirement of several introductory psychology courses. Each of the subjects took part in both studies. Subjects worked individually in the presence of the experimenter. They were not paid for their participation, but were able to actually play three of the lotteries they viewed, receiving their winnings. The maximum payoff for the three lotteries was $30, and the minimum payoff was $0.

Lotteries

A general representation of an ambiguous lottery appears in Fig. 2. There is an imprecise probability p of receiving the outcome x, and a complementary imprecise probability of receiving nothing. The probability of the outcome x may lie anywhere within the interval $[P_{min}, P_{max}]$, having range $R = P_{max} - P_{min}$ and center P. Three parameters specify the lottery: the center of the interval P, the outcome x, and the range of imprecision R. In general, whenever $R > 0$, the lottery is ambiguous; and when $R = 0$, the lottery is nonambiguous. All the lotteries in Studies 1 and 2 were of this form.

Procedure

The experiment involved a number of lotteries among which the subject indicated his or her preferences by ordering a set of cards, each of which contained a lottery display. Each lottery had two possible monetary outcomes. The outcomes depended upon the draw of a poker chip from a bag containing 100 chips. All lotteries involved a common five-step playing procedure. This procedure, and the displays used to represent the lotteries, were described by written instructions. One of the displays used to demonstrate the procedure was Lottery 2, shown in Fig. 3. The description of the steps in playing this lottery, as presented in the instructions, follows:

1. Receive a bag containing 60 chips. Each of the chips in the bag is either BLUE or WHITE. You do not know how many of the 60 chips are BLUE, or how many are WHITE. You know only the total number of chips in the bag, and that each chip is either BLUE or WHITE. *You cannot check the contents of this bag.* It is these 60 unknown chips that are represented by the black region of the bar graph.

2. Designate either BLUE or WHITE as your WINNING color for this lottery. Note that the bag is in your possession (even though you cannot yet check

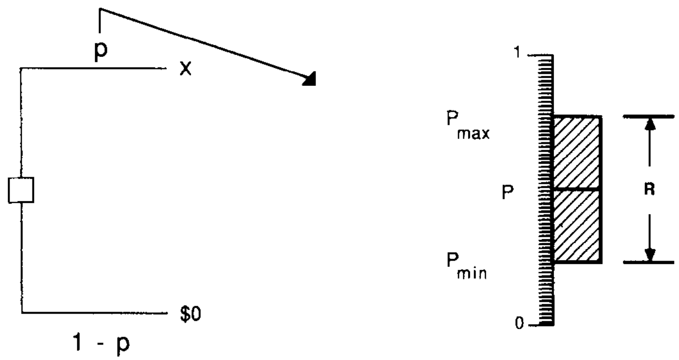

FIG. 2. Representation of an ambiguous lottery and the notation used to describe the features of the lottery.

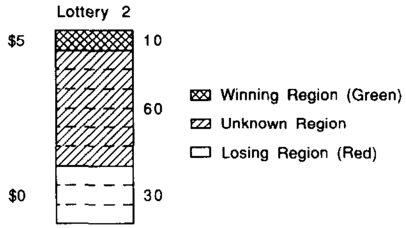

FIG. 3. Display of Lottery 2, a sample lottery used in the experiment. A green region represents 10 winning ($5) chips, a red region represents 30 losing ($0) chips, and 60 chips are of unknown color and not known to be either winning or losing.

its contents), and therefore the bag's contents cannot be biased either for or against you after you designate your color.

3. Place 10 WINNING chips [green region] and 30 LOSING chips [red region] in the bag already containing 60 chips. Thus, the bag contains 10 chips that you know are WINNING chips, 30 chips that you know are LOSING chips, and 60 chips that may be any combination of WINNING and/or LOSING chips.

4. Draw a chip without looking. If the chip is a WINNING chip, you receive the upper amount pictured in Lottery 2 of $5; if the chip is a LOSING chip, you receive the lower amount shown of $0.

5. At this point, you may check that the bag contains what it is said to contain. If the bag does not contain what it should and you did not win, then the lottery is repeated.

The instructions also described the experimental task and the procedure by which the three lotteries would be picked for the subject to play.

Following the instructions, the subject's understanding of the lottery displays was tested by having the subject describe to the experimenter the playing procedure for two sample lottery displays. The subject was then reminded: (a) that the task was to rank, in order of preference, sets of lottery display cards; (b) that all the lotteries involved 100 total chips, though the numbers of WINNING, LOSING, and unknown chips varied; (c) that all lotteries had two possible outcomes, though the outcomes varied between lotteries; and (d) that three of the lotteries would be played at the end of the session, that the selected lotteries would partly depend on their preferences, and that the subject would receive the sum total of the winnings from the three lotteries.

Regarding this latter point, it was noted that some of the lotteries involved losses, and that the total winnings of a subject could be negative. In such an event, they received $0. However, losses were real to the subject in the sense that a loss in one lottery could cancel a win that the subject would otherwise have received from another of the three lotteries played.

After any questions, subjects participated in Studies 1 and 2 of the experiment, in that order. The details of the procedures of these studies are presented below conjoint analysis is its ability to reject polynomial models, an ability which can greatly advance the present goal of narrowing the set of possible models.

At the same time, conjoint measurement has a major weakness. The theory basically does not accommodate any errors. Even one violation of a property, in principle, is to be interpreted as a failure of the property and any model for which the property is necessary. In applying the theory to empirical data, a softer stand, one which allows errors, must be developed. To do this, a number of error theory proposals that may be useful for particular applications have been advanced (cf. Weber, 1984, for a review of various proposals). However, there is no currently accepted approach for handling errors.

Diagnostic Properties

The necessary conditions of diagnostic importance for the polynomials in Table 1 are independence conditions. The properties are defined in this section for three factors, labelled X, Y, and Z. The relation $\geqslant_p$ is used to designate the preference relation "is at least as preferred as."

A factor X is said to be *independent* of a factor Y, holding a factor Z constant, if and only if, for all x_1, x_2 in X, z_1 in Z,

$$x_1 y_1 z_1 \geqslant_p x_2 y_1 z_1$$

for some y_1 implies that

$$x_1 y_2 z_1 \geqslant_p x_2 y_2 z_1$$

for every y_2 in Y. Thus, X is independent of Y holding Z constant if the ordering induced on X with Y and Z fixed is independent of the level of Y. This independence relation is summarized as $X;\ Y:Z$, and is illustrated by Fig. 8. The ordering represented by each of the three line segments in the figure must be the same to support the conclusion that the property of independence $X;\ Y:Z$ holds. This must be true for each of the planes in the figure defined by the level at which Z is fixed. However, the ordering over X does not need to be the same for all levels of Z. For example, the independence property $X;\ Y:Z$ implies that if the option designated by Cell a (Fig. 8) is preferred to the option designated by Cell b, then Cell a′ must be preferred to Cell b′; it is not implied that Cell c must be preferred to Cell d.

For simple polynomial models, the relation $X;\ Y:Z$ is a necessary condition for factors X and Y which combine additively, $X+Y$, or for factors X and Y which combine multiplicatively, $X*Y$, with Y assuming only positive values. For three factors, there are six possible independence properties: $X;\ Y:Z$, $X;Z:Y$, $Y;X:Z$, $Y;Z:X$, $Z;X:Y$, and $Z;\ Y:X$.

For the related condition of sign dependence, the set of levels of the factor Y is partitioned into three sets: Y^+, Y^0, and Y^-. A factor X is said to be *sign dependent* on Y, holding Z constant, if and only if Y can be so partitioned with:

Procedure

Within the general experimental procedure, already presented, the following was the specific procedure for Study 1. Subjects were given four lotteries at a time and were asked to order these according to preference. One of the four-lottery sets, Set 1–1, is displayed in Fig. 5. Each set contained an ambiguous Lottery A_1 having a winning outcome x_1 (\$10.01), an ambiguous Lottery A_2 having a smaller winning outcome x_2 (\$10), a nonambiguous Lottery N_1 having winning outcome x_1, and a nonambiguous Lottery N_2 having winning outcome x_2. For all ambiguous lotteries A in Study 1, the number of unknown chips was 40. For all nonambiguous lotteries N, the number of unknown chips was 0. The losing outcome was \$0 for all lotteries.

For ambiguous options, we interpret the center of the unknown probability interval as the best-guess probability for the nonzero outcome x. Thus, the best-guess probability $P = .50$ for all the lotteries in Fig. 5. This is primarily by assumption, although subjects, when asked, agreed with this assumption. Subjects uniformly believed that blue and white were equally likely in the composition of the bags of unknown chips in all Lotteries A_1 and A_2 used in the study.

Clearly, Lottery N_1 dominates N_2 in each lottery set; and similarly, Lottery A_1 dominates A_2. These dominance relations were required of each subject's rankings. Under these dominance constraints, six rankings of the four lotteries in each set were possible (from most to least preferred within each set):

1. N_1, N_2, A_1, A_2;
2. A_1, A_2, N_1, N_2;
3. N_1, A_1, N_2, A_2;

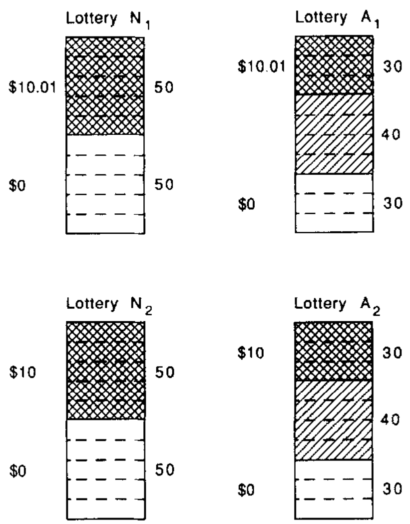

FIG. 5. Displays of Lotteries N_1, A_1, N_2, and A_2 included in Set 1–1 of Study 1.

4. A_1, N_1, A_2, N_2;
5. N_1, A_1, A_2, N_2;
6. A_1, N_1, N_2, A_2.

If a subject used a lexicographic decision rule, then he or she would respond with either Ranking 3 or Ranking 4. Ranking 3 signifies a lexicographic form of ambiguity avoidance; Ranking 4 signifies a lexicographic form of ambiguity seeking. If a subject was trading off expectation and ambiguity, then he or she would respond with either Ranking 1 or Ranking 2. Ranking 1 signifies non-lexicographic ambiguity avoidance; Ranking 2 signifies non-lexicographic ambiguity seeking. Rankings 5 and 6 are consistent with neither a tradeoff nor a lexicographic model, and were used for comparison.

Eight series of four-lottery sets were presented to each subject. The eight series are summarized by Table 2 and detailed in Table 3. First, consider Series 1, which is outlined in the second columns of Tables 2 and 3. This series involves systematically increasing x, the winning amount, upward from \$10 for Lotteries A_1 and N_1. Suppose the subject gave Ranking 1 for the first lottery in this set, Set 1–1 (the set illustrated by Fig. 5). The subject was characterized as willing to trade off \$.01 to avoid ambiguity in the vicinity of $P = .50$. This was concluded from the preference for Lottery N_2 over Lottery A_1. The subject then received Set 1–2 to determine if he or she would be willing to trade off \$.05, and so on, until a point was reached

TABLE 2

Summary of the Eight Series of Four-Lottery Sets (Study 1)

	Series i							
	1	2	3	4	5	6	7	8
Winning Amount of A_1, N_1	varies \$10.01–\$14	\$10	\$10	\$10	varies \$10.01–\$18	\$10	\$10	\$10
Winning Probability P for A_1, N_1	.50	varies .51–.70	.50	.50	.25	varies .26–.45	.75	.75
Winning Amount of A_2, N_2	\$10	\$10	varies \$9.99–\$6	\$10	\$10	\$10	varies \$9.99–\$7.33	\$10
Winning Probability P for A_2, N_2	.50	.50	.50	varies .49–.30	.25	.25	.75	varies .74–.55

Note: All lotteries were two-outcome lotteries, with losing amount \$0, involving one draw from a bag containing 100 poker chips. Ambiguous lotteries A_1 and A_2 had 40 unknown chips; nonambiguous lotteries N_1 and N_2 had 0 unknown chips.

TABLE 3

Values of the Varying Parameter and Manipulations Used in Each of the Eight Lottery Series (Study 1)

	Series *i* with Manipulations							
Set *i*–*n*	1 $x\uparrow$ near $P=.50$	2 $P\uparrow$ near $P=.50$	3 $x\downarrow$ near $P=.50$	4 $P\downarrow$ near $P=.50$	5 $x\uparrow$ near $P=.25$	6 $P\uparrow$ near $P=.25$	7 $x\downarrow$ near $P=.75$	8 $P\downarrow$ near $P=.75$
i–1	$10.01	—	$9.99	—	$10.01	—	$9.99	—
i–2	$10.05	—	$9.95	—	$10.05	—	$9.95	—
i–3	$10.20	.51	$9.80	.49	$10.40	.26	$9.87	.74
i–4	$10.60	.53	$9.40	.47	$11.20	.28	$9.60	.72
i–5	$11.00	.55	$9.00	.45	$12.00	.30	$9.33	.70
i–6	$12.00	.60	$8.00	.40	$14.00	.35	$8.67	.65
i–7	$13.00	.65	$7.00	.35	$16.00	.40	$8.00	.60
i–8	$14.00	.70	$6.00	.30	$18.00	.45	$7.33	.55

Note: In describing the manipulations, P represents the winning probability of the lotteries in Set i for which that probability varies. Similarly, x represents the winning amount of the lotteries in Set i for which that amount varies. The arrows indicate the direction in which the probability or amount was altered, and the neighboring probability of winning for each series is also shown. For each series, subjects saw Sets i–n in order of increasing n.

at which the subject no longer responded with Ranking 1 for this series. In this way, using a choice paradigm, an approximate measure of the subject's strength of reaction to ambiguity in the vicinity of $P = .50$ was obtained by systematically increasing x.

Similarly, Series 2 provided an approximate measure of the strength of ambiguity reactions in the vicinity of $P = .50$ by systematically increasing P, Series 3 in the vicinity of $P = .50$ by decreasing x, Series 4 in the vicinity of $P = .50$ by decreasing P, Series 5 in the vicinity of $P = .25$ by increasing x, Series 6 in the vicinity of $P = .25$ by increasing P, Series 7 in the vicinity of $P = .75$ by decreasing x, and Series 8 in the vicinity of $P = .75$ by decreasing P. These manipulations are summarized in the top row of Table 3, with the values that were used for the varying quantity in each series being listed in the body of the table.

Decreasing values of P and x were not used in the vicinity of $P = .25$ because of the "floor" of $P = 0$ that does not allow sufficient variation in P. Similarly, $P = 1$ acts as a "ceiling" which prohibits increasing P in the vicinity of $P = .75$. In the vicinity of $P = .50$, no "ceiling" or "floor" operates. Both increasing and decreasing P and x series were included in this region so as to compare the results at $P = .50$ with the results at $P = .25$ and $P = .75$, respectively.

Series 1 and 2 were presented concurrently. Both involved Lotteries A_2 and N_2, having best-guess probability $P = .50$ of receiving the nonzero outcome $x = \$10$. The other two lotteries in each set, A_1 and N_1, were varied. The subject received Set 1–1, and ranked the four lottery cards according to preference. The subject then received Set 2–3 and ranked that set. The subject continued receiving cards from alternate sets, moving down through each series, as long as either Ranking 1 or Ranking 2 was obtained. Once a ranking from Rankings 3 through 6 was given, that series was ended. Thus, for each subject, for each series, the crossover point between Rankings 1 and 2 and Rankings 3 through 6 was determined.

Series 3 and 4, like Series 1 and 2, were presented concurrently, as were Series 5 and 6 and Series 7 and 8. The order of presentation of these series pairs was randomized across subjects.

Measures

For Lottery Set i–n within Series i, let $E_1(i\text{–}n)$ designate the expectation of Lotteries A_1 and N_1, and let $E_2(i\text{–}n)$ designate the expectation of Lotteries A_2 and N_2, using the best-guess probabilities for the lotteries in calculating these expectations. Then, for Series 1 and 2, $E_2(1\text{–}n) = E_2(2\text{–}n) = \$10 * .50 = \$5$ for all the lottery sets, and $E_1(i\text{–}n)$ increases as each of these two series progresses (as n increases). The lotteries in Series 1 and 2 increase in a systematic fashion, with $E_1(1\text{–}n) = E_1(2\text{–}n)$, for $n \geqslant 3$. Further, since $E_2(1\text{–}n) = E_2(2\text{–}n) = \$5$, for all n, it also holds that

$$E_1(1\text{–}n) - E_2(1\text{–}n) = E_1(2\text{–}n) - E_2(2\text{–}n).$$

In fact, all eight series in Tables 2 and 3 are designed such that

$$E_1(i\text{–}n) - E_2(i\text{–}n) = E_1(j\text{–}n) - E_2(j\text{–}n) \qquad (1)$$

for all Series i and j in Study 1 for $n \geqslant 3$. This also holds approximately for $n = 1$ and $n = 2$. Since Eq. (1) holds for all the series, a measure of ambiguity reactions which is comparable across series is specifiable. The *standardized ambiguity-avoidance measure* for Series i is defined by

$$\mathrm{SA}(i) = \begin{cases} E_1(i\text{–}N) - E_2(i\text{–}N), & \text{for } i\text{–}N \text{ having Ranking 1} \\ E_2(i\text{–}N) - E_1(i\text{–}N), & \text{for } i\text{–}N \text{ having Ranking 2} \\ 0, & \text{for } i\text{–}1 \text{ having Ranking 3 through 6,} \end{cases}$$

where i–N is the last set in Series i having a ranking of 1 or 2. Thus, SA(i) is positive for ambiguity-avoiders who respond with Ranking 1 to Set i–1, negative for ambiguity-seekers who respond with Ranking 2 to Set i–1, and 0 for those who do not trade off to avoid or seek ambiguity, responding with Rankings 3 through 6 to Set i–1.

For example, suppose a subject gave Ranking 1 for Sets 1–1 through 1–3, and then responded with Ranking 5 for Set 1–4. For this subject and this series, $N = 3$; thus, \$.20 was traded off to avoid ambiguity, but \$.60 was not. The series would be ended at this point. The subject was willing to trade off an expectation $\mathrm{SA}(1) = E_1(1\text{–}3) - E_2(1\text{–}3) = \$5.10 - \$5.00 = \$.10$. Alternatively, suppose a subject responded with Ranking 2 for Sets 2–3 through 2–5, and then responded with Ranking 4 for Set 2–6, having $N = 5$. Then, $\mathrm{SA}(2) = E_2(2\text{–}5) - E_1(2\text{–}5) =$

\$5.00 – \$5.50 = –\$.50. The subject was willing to trade off \$.50 in expectation to play a more ambiguous lottery, the negative sign signifying ambiguity seeking.

For each Series *i* for each subject, the strength of ambiguity avoidance was measured by SA(*i*). The measure is a conservative indicator of ambiguity reactions, being a lower bound for the magnitude of a subject's reaction to ambiguity, whether positive or negative. The exemplar subject of the previous paragraph with SA(1) = \$.10, for instance, might actually have been willing to trade off more than \$.20 to avoid ambiguity in Series 1. The subject's maximal tradeoff may have been anywhere within the interval [\$.20, \$.60), implying that SA(1) could lie within [\$.10, \$.30). The lower quantity, in absolute value, was taken as our estimate of the subject's strength of ambiguity avoidance, and the definition of SA(*i*) stated appropriately.

Results and Discussion

Regarding the validity of the lexicographic model, Model 2 in Table 1, each subject's response to the first set within each series was noted. The response frequencies are displayed in Fig. 6. Using the frequencies of Rankings 5 and 6 for these first sets as error rates, Rankings 1 and 2 were significantly more prevalent (two-tailed binomial, $p < .001$ for all eight tests). The frequencies of Rankings 3 and 4 did

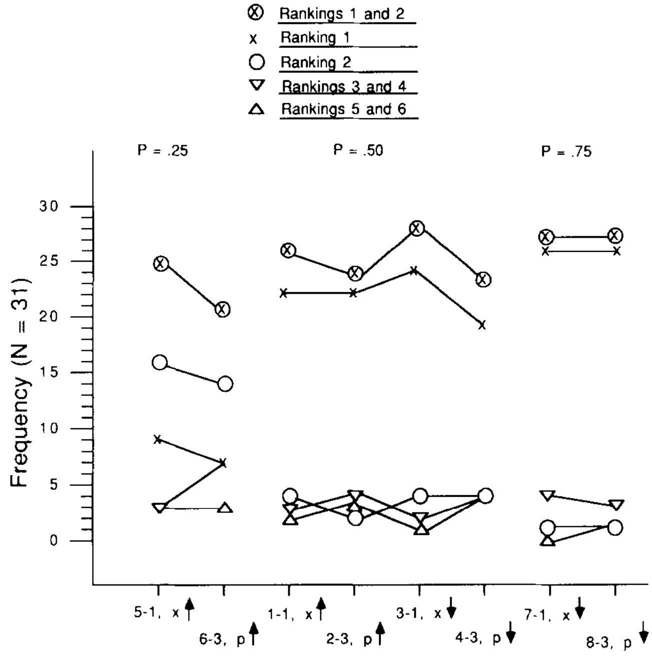

FIG. 6. Frequencies of Rankings 1 through 6 as responses in each of the first sets in Series 1 through 8 of Study 1. The manipulation used for each series and the neighboring probability *P* are indicated.

not significantly exceed those of Rankings 5 and 6 ($p > .1$ for all tests). Further, the frequencies of Rankings 1 and 2 exceeded those of Rankings 3 and 4 ($p < .02$ for all tests).

This consistent pattern of results across levels of P and manipulations implies that subjects were indeed willing to trade off expectation in their reactions to ambiguity. The lexicographic response patterns, Rankings 3 and 4, did not exceed a chance rate of occurrence. Thus, the lexicographic model was clearly inadequate as descriptive of subjects' choices.

Comparing the frequency of Ranking 1 to that of Ranking 2, an effect of the level of the best-guess probability was observed. Ranking 1 predominated near $P = .50$ and $P = .75$ ($p < .02$ for all tests), with Ranking 2 not exceeding the frequencies of Rankings 5 and 6 ($p > .3$). However, near $P = .25$, the frequency of Ranking 1 did not significantly differ from those of Rankings 5 and 6 ($p > .1$), but the frequency of Ranking 2 did ($p < .02$). Moreover, the frequencies of Rankings 1 and 2 did not significantly differ ($p > .1$).

Since, at $P = .25$, there was a significant reaction to ambiguity, in that the frequencies of Rankings 1 and 2 exceeded a chance rate, and since Ranking 1 did not predominate, a significant level of ambiguity-seeking behavior was identified. This result, in contrast with that reported by Curley and Yates (1985), is most likely attributable to the greater sensitivity of the present analysis. The lack of a significant difference between the frequencies of Rankings 1 and 2 was comparable to the test for ambiguity seeking used in the Curley and Yates study, which failed to identify the behavior. In agreement with that study and others, ambiguity avoidance predominated at $P = .50$ and $P = .75$. Interestingly, there was no indication of ambiguity seeking at either of these probability levels.

The positive relationship between ambiguity avoidance and the probability level P is also evidenced by Fig. 7, which shows the mean standardized ambiguity-avoidance measure SA(i) for each Series i. The 95% confidence interval for each mean measure is also indicated by the figure. Comparison of series across levels of P, involving either changes in P or x, reveals that the mean SA measure differs between all pairs of P values, $P = .25$, .50, and .75 (two-tailed $t(30)$, $p < .05$ for all tests).

The different within-P manipulations did not differ at any of the three levels of P ($p > .05$ for all tests). Thus, there is some consistency in the mean SA measures obtained. In the neighborhood of $P = .75$, subjects, on average, were willing to give up an expectation of \$.65 to avoid the ambiguous option. Near $P = .50$, subjects would give up an expectation of approximately \$.31 to avoid the ambiguous option. Near $P = .25$, subjects were willing to trade off an expectation of \$.14 to obtain the *more* ambiguous option over the nonambiguous one.

The value of \$.31 near $P = .50$ is comparable to the value obtained by Yates and Zukowski (1976), who used a pricing task and slightly different lotteries. That study and the present study both obtained values greater than those reported by Becker and Brownson (1964) in the region of $P = .50$. The latter study also used a pricing procedure, but did not encourage honest reporting of subjects' values.

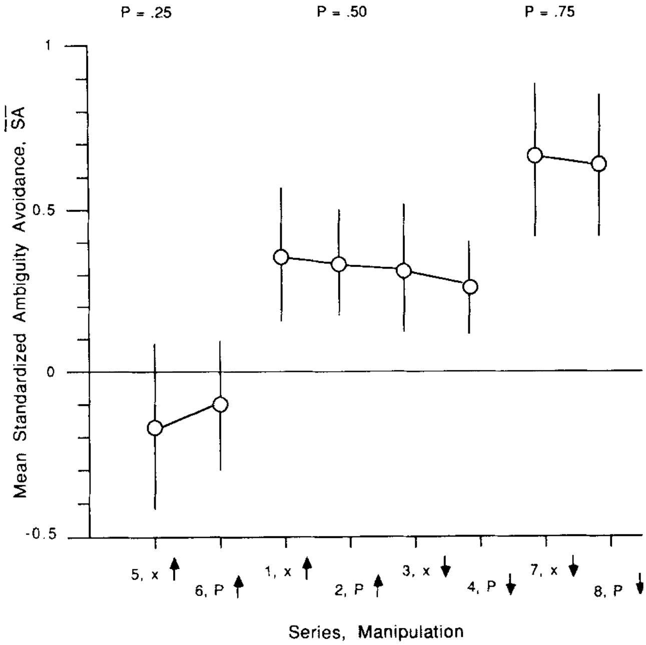

FIG. 7. Mean standardized ambiguity-avoidance measure $\overline{SA}$ for each of Series 1 through 8 used in Study 1, along with the 95% confidence interval for each mean value. The manipulation used for each series is indicated.

Study 2

Study 1 indicated that the lexicographic model, Model 2, could be eliminated as a viable model of choice under ambiguity. Model 1 has also been rejected, in that it does not account for the reactions to ambiguity that subjects have displayed. Study 2 was designed primarily to evaluate the remaining models in Table 1, the three-factor models, Models 3–6, but also provided evidence for evaluating the adaptation of portfolio theory that was described. The study examined the polynomial models by means of analyses based on the theory of polynomial conjoint measurement, which originated with Luce and Tukey (1964). The conjoint-analytic approach to differentiating polynomial models involves a series of necessary properties which are tested as part of a diagnostic procedure (Krantz & Tversky, 1971). The procedure that was used involved several modifications to the analytic process and tests described by Krantz, Luce, Suppes, and Tversky (1971), and is described in more detail by Curley (1986, 1989).

One of the clear benefits of conjoint analysis is that it makes minimal scale assumptions about the data. In particular, only the ordering of options with respect to the undefined ambiguity factor is required for differentiation among the models. As such, conjoint analysis is a very robust procedure. A second attraction of conjoint analysis is its ability to reject polynomial models, an ability which can greatly advance the present goal of narrowing the set of possible models.

At the same time, conjoint measurement has a major weakness. The theory basically does not accommodate any errors. Even one violation of a property, in principle, is to be interpreted as a failure of the property and any model for which the property is necessary. In applying the theory to empirical data, a softer stand, one which allows errors, must be developed. To do this, a number of error theory proposals that may be useful for particular applications have been advanced (cf. Weber, 1984, for a review of various proposals). However, there is no currently accepted approach for handling errors.

Diagnostic Properties

The necessary conditions of diagnostic importance for the polynomials in Table 1 are independence conditions. The properties are defined in this section for three factors, labelled X, Y, and Z. The relation $\geqslant_p$ is used to designate the preference relation "is at least as preferred as."

A factor X is said to be *independent* of a factor Y, holding a factor Z constant, if and only if, for all x_1, x_2 in X, z_1 in Z,

$$x_1 y_1 z_1 \geqslant_p x_2 y_1 z_1$$

for some y_1 implies that

$$x_1 y_2 z_1 \geqslant_p x_2 y_2 z_1$$

for every y_2 in Y. Thus, X is independent of Y holding Z constant if the ordering induced on X with Y and Z fixed is independent of the level of Y. This independence relation is summarized as X; $Y:Z$, and is illustrated by Fig. 8. The ordering represented by each of the three line segments in the figure must be the same to support the conclusion that the property of independence X; $Y:Z$ holds. This must be true for each of the planes in the figure defined by the level at which Z is fixed. However, the ordering over X does not need to be the same for all levels of Z. For example, the independence property X; $Y:Z$ implies that if the option designated by Cell a (Fig. 8) is preferred to the option designated by Cell b, then Cell a′ must be preferred to Cell b′; it is not implied that Cell c must be preferred to Cell d.

For simple polynomial models, the relation X; $Y:Z$ is a necessary condition for factors X and Y which combine additively, $X+Y$, or for factors X and Y which combine multiplicatively, $X*Y$, with Y assuming only positive values. For three factors, there are six possible independence properties: X; $Y:Z$, X; $Z:Y$, Y; $X:Z$, Y; $Z:X$, Z; $X:Y$, and Z; $Y:X$.

For the related condition of sign dependence, the set of levels of the factor Y is partitioned into three sets: Y^+, Y^0, and Y^-. A factor X is said to be *sign dependent* on Y, holding Z constant, if and only if Y can be so partitioned with:

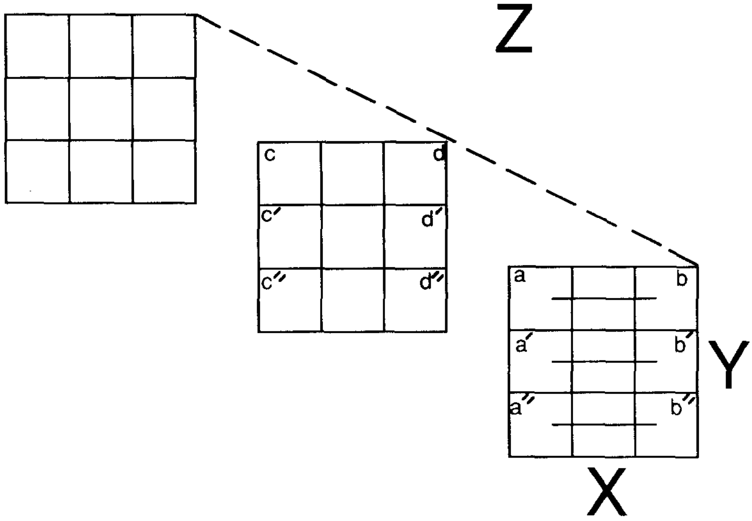

FIG. 8. Diagram demonstrating the independence property X; $Y : Z$. Shown are three levels of each of the three variables X, Y, and Z. Line segments connect those cells whose orderings are to be compared. Letters denote exemplar cells in the diagram: a, a', a'', b, b', etc.

(i) independence holding over each of Y^+ and Y^- separately, that is, X; $Y^+ : Z$ and X; $Y^- : Z$ hold;

(ii) any ordering induced on X with Y^+ and Z fixed being the reverse of any ordering induced on X with Y^- and Z fixed;

(iii) the ordering over X with Y^0 and Z fixed being degenerate.

The sign dependence relation is summarized as X; $Y^* : Z$.

In Fig. 8, any nondegenerate ordering over the cells connected by a line segment would have to be either identical to or the exact reverse of any ordering over the cells connected by other line segments. The direction of the orderings separates the levels of Y into the two categories Y^+ and Y^-, with indifference over all the cells connected by a particular line segment characterizing levels in the category Y^0.

Note that sign dependence does impose a requirement across levels of Z. For example, suppose Cell a is preferred to Cell b (Fig. 8). We can arbitrarily label the level of Y for these cells as being within the category Y^+. Further suppose Cell b′ is preferred to Cell a′. This implies that the level of Y for these cells is necessarily within the category Y^-. Now, if Cell c is preferred to Cell d (this is not a necessary implication of X; $Y^* : Z$) when the level of Y is within Y^+, then it is implied that Cell d′ must be preferred to Cell c′ when the level of Y is within Y^-. Similarly, if the subject is indifferent between Cell a″ and Cell b″, indicating that the level of Y for these cells is within Y^0, then the property implies that the subject should be indifferent between Cell c″ and Cell d″.

For simple polynomial models, the relation X; $Y^* : Z$ is a necessary condition for factors X and Y which combine multiplicatively, $X * Y$, with Y assuming positive and nonpositive values. Similarly to independence, there are six properties of sign dependence possible for the three-factor case. Which of the two properties, $X; Y : Z$ or $X; Y^* : Z$, is appropriate depends upon the structure of Y.

Other diagnostic properties which are available, but which were not important in Study 2, are the joint independence properties and various cancellation conditions. Their use in practical applications of the conjoint measurement axioms is described by Curley (1989), and their use in Study 2 is presented by Curley (1986).

In applying the tests implied by conjoint measurement theory within Study 2, the three factors comprised in the models—A, P, and U—were varied. Operationally, lotteries of the form described by Fig. 2 were employed which differed in the nonzero outcome x (factor U), the central probability P of the nonzero outcome (factor P), and the uncertainty R about this probability (factor A). The analytical measures used in Study 2 are described along with the results.

Procedure

Within the general experimental procedure and following the procedure for Study 1, both already described, the following was the procedure specific to Study 2. Two-outcome lotteries were presented on cards, using displays like that in Fig. 3. For all lotteries, \$0 was one of the outcomes; the nonzero outcome x was varied. All lotteries involved 100 chips, with the number of winning, losing, and unknown chips varying among lotteries. Subjects were reminded of these lottery characteristics, and then informed of the experimental task.

Subjects ranked sets of ten lotteries according to their preference among the lotteries, from most to least preferred. To aid in this task, the following method of accumulation was suggested (Coombs & Bowen, 1971). The subject begins with two lotteries from the set and decides which is preferred. The card with the preferred lottery is placed to the left of the other. The remaining eight lotteries, one at a time, are then placed relative to those already ordered. Finally, the subject checks the final ordering for any corrections. This method was suggested, but subjects were advised to use whatever method they felt appropriate, as long as the ordering was obtained.

Nineteen lotteries were used in the study, allowing the use of a balanced incomplete block design. Nineteen sets of lottery cards, each consisting of 10 lotteries, were ranked by each subject. The order of the sets was randomized across subjects. The design has an efficiency of 95%, and provided five replications of each lottery pair (Cochran & Cox, 1957).

Eighteen of the lotteries were the factorial combination of two or three levels on three factors: U (2 levels), P (3 levels), and A (3 levels). The levels used were:

U: $x = \$10$ and $-\$10$

P: $P = .25$, $.50$, and $.75$

A: $R = 0$, $.20$, and $.50$,

where x was the nonzero outcome with best-guess probability P of occurring and R was the range within which the probability p was uncertain. For example, the displays for the three lotteries having $x = -\$10$ and $P = .50$ are illustrated by Fig. 9. The levels of P and R were chosen so as to span the sets of possible values. The values of x were selected to obtain tests of sign dependence, particularly of A, upon the factor U. In addition to these eighteen, the lottery having $x = \$5$, $P = .50$, and $R = .50$ was added. This particular lottery was included to provide insight into the application of portfolio theory as a choice model under ambiguity, as described with the results.

Results and Discussion

Consistency. The five replications of each lottery pair allowed a determination of the consistency of subjects' responses. Note that, of the 19 lotteries in the design, 10 offered either a gain or nothing and nine promised either a loss or nothing. Clearly, the former dominated the latter, and this dominance should be reflected in a subject's rankings. In addition, for the 18 lotteries in the factorial design, there was a probability dominance present. That is, within the 9 lotteries involving a gain or nothing, those having $P = .75$ dominated those having $P = .50$, which in turn dominated those having $P = .25$. This was true regardless of the level of R, for the values of R included in the study. The 18 lotteries in the factorial design were thereby partitioned into 6 groups of three lotteries each, as defined by the levels of P and x. One of these groups is shown by Fig. 9.

Intrasubject consistency was measured within these groups by Kendall's (1970) coefficient of agreement u. In the present design, values ranged between $u = -.20$, indicating minimal consistency, and $u = 1.0$, indicating maximal consistency. A coefficient was determined for each subject, for each of the 6 groups of lotteries, separately; its values are listed in Table 4. The coefficients indicate considerable intraindividual consistency. For example, 71.0% of the coefficients were at or above

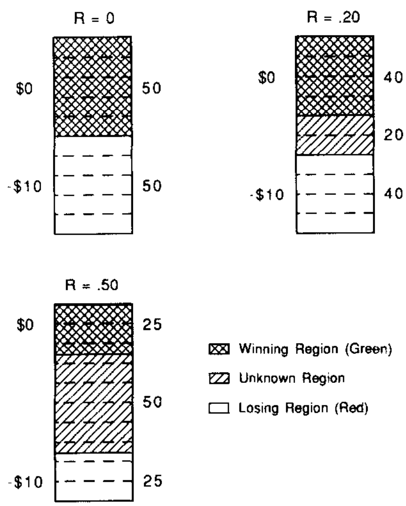

FIG. 9. Displays of the three lotteries in Study 2 having $x = -\$10$ and $P = .50$.

TABLE 4

Frequencies of the Coefficients of Agreement, Study 2

u	f	Significance level[a]
−.20	9	
−.07	8	
.07	13	
.20	11	
.33	13	
.47	18	$p < .05$
.60	16	
.73	34	$p < .01$
1.00	64	$p < .001$

[a] Statistical significance of values of u this extreme, on the assumption that responses were allotted purely at random.

the value $u = .47$, with $p < .05$, in comparison with the model stating that subjects behaved randomly. Thus, the reliability of the subjects' responses was accepted, in that there was consistency of their responses among the replications.

Preference Ordering. In order to apply the independence properties, the 5 replications for each lottery pair were compressed into a dominant preference ordering. This was done by first creating a 19×19 preference matrix for each subject. The entry in cell (i, j) of the matrix ranged from 0 to 5 and indicated the number of times that Lottery i was preferred to Lottery j. The entries were then summed over the columns, and the ordering of the row sums used as an ordering of the 19 stimuli for that subject. This method had the advantage of directly using all the information provided in the data, although an alternative method for producing the dominant ordering was also employed and led to equivalent results (Curley, 1986).

Independence and Sign Dependence. For the present study, the six possible properties of independence or sign dependence were not equally diagnostic. In particular, certain of the independence conditions were expected solely by virtue of the dominance relations that have been indicated. Specifically, it was expected that the independence properties $U; A : P$, $U; P : A$, $P; A : U$, and the sign dependence $P; U^* : A$ would obtain.

The data supported these expectations. As summarized in Table 5, the number of subjects who perfectly satisfied these properties far exceeded a chance rate. Overall, 24/31 (77.4%) of the subjects satisfied all four of these conditions perfectly.

Of greater interest were the remaining two independence properties, $A; P : U$ and $A; U : P$. First, note that the independence properties involve the extent to which two or more preference rankings correspond. As such, the Spearman rank correlation coefficient, r_s, was used in summarizing the data. The coefficient assumes the value 1.0 when two rankings perfectly correspond, consistent with independence

TABLE 5

Number of Subjects Perfectly Satisfying the Dominance-Implied Properties of Independence and Sign Dependence

Independence/ Sign Dependence Property	# Satisfying the Property ($N=31$)	Expected # under Random Model
$U; A : P$	27	.48
$U; P : A$	26	.48
$P; A : U$	29	.03
$P; U^* : A$	29	.15

and sign dependence within the same sign category; and it assumes the value -1.0 when two rankings are perfectly opposed, consistent with sign dependence between sign categories.

For each subject for each property, values of r_s were calculated between each pair of orderings relevant for the property. The mean of the appropriate values of r_s was used as an indicator of the degree to which the property obtained. For the property $A; P : U$, there were three rankings over A at each fixed level of x. A value of r_s was determined within each of the three possible pairs which could be drawn from this triple of rankings. Thus, each subject's mean rank correlation $\bar{r}_s$ was based upon six values of r_s, there being three interdependent values of r_s for each of two fixed levels of x. Perfect sign dependence $A; P^* : U$ was indicated by an $\bar{r}_s = -.33$, and perfect independence $A; P : U$ was indicated by an $\bar{r}_s = 1.0$. For the property $A; U : P$, each subject's mean rank correlation $\bar{r}_s$ was based upon three values of r_s, one at each fixed level of P. Perfect sign dependence $A; U^* : P$ was indicated by an $\bar{r}_s = -1.0$, and perfect independence $A; U : P$ by an $\bar{r}_s = 1.0$. To allow for some error in subjects' responding, perfect satisfaction of the properties was not required in the analysis. A subject whose orderings were within one pairwise reversal of either perfect independence or sign dependence was categorized as adequately satisfying the corresponding property.

The two properties are not independent of each other; so a summary of the joint distribution is illustrated by Table 6. The frequencies in Table 6 indicate that nine subjects were within one pairwise reversal of perfect independence for both properties, with values of $\bar{r}_s \geqslant .83$. Since only .05 subjects would be expected to be classified in this cell of the table if they were responding randomly, the data support the existence of a group of subjects who satisfied both independence properties $A; P : U$ and $A; U : P$.

There is a second, minor clustering of subjects having $\bar{r}_s$ within $[-.50, -.17]$ for the property $A; P : U$ and $\bar{r}_s$ within $[-.67, -.17]$ for the property $A; U : P$. These subjects can be described as satisfying sign dependence $A; P^* : U$ within one pairwise reversal, but neither $A; U^* : P$ nor $A; U : P$. A better characterization of these subjects is possible, however, as is now described.

TABLE 6

Observed Frequencies and Expected Frequencies under Random Model of Values $\bar{r}_s$ for Properties $A; U : P$ and $A; P : U$

	$A; U : P$			
$A; P : U$	$[-1, -.83]$[a]	$[-.67, -.17]$	$[0, .67]$	$[.83, 1]$[b]
		Actual Frequencies		
$[-.50, -.17]$[a]	0	9	1	0
$[-.08, .67]$	0	2	8	2
$[.83, 1]$[b]	0	0	0	9
		Expected Frequencies		
$[-.50, -.17]$[a]	.50	5.23	6.13	.50
$[-.08, .67]$	.46	7.44	9.98	.46
$[.83, 1]$[b]	.05	.10	.10	.05

[a] Values consistent with sign dependence version of this property within one pairwise inversion.
[b] Values consistent with independence version of this property within one pairwise inversion.

The values $x = -\$10$ and $x = \$10$ were selected to manipulate the factor U as a signed factor. It was intended that subjects would perceive the stationary outcome \$0 as a reference point of zero utility against which $x = -\$10$ would have negative value and $x = \$10$ would have positive value (along the lines of Kahneman & Tversky, 1979). The frequencies in Table 6 are reported under this assumption; however, the assumption is not a necessary requirement of the models. Suppose the assumption were not valid. Suppose, instead, that subjects used $-\$10$, or a smaller amount, as the reference point, against which none of the outcomes would have negative utility. Further, the probability P of the nonzero outcome would no longer be the relevant probability. Instead, P', defined as the probability of the greater outcome, whether \$10 or \$0, may be a more accurate index of the factor P. When $x = -\$10$, $P = .25$ implies $P' = .75$ and $P = .75$ implies $P' = .25$; otherwise $P = P'$ in the design. Altering this assumption only affects the results for the test of the independence property $A; U : P$. The independence property under the assumption of no negative utilities is designated as $A; U : P'$ to distinguish it from the property $A; U : P$ assuming negative utilities and reported in Table 6.

Losing the assumption of \$0 as a reference point apparently reduced the capability of the study to test the models of concern, in that the signedness of the factor U was not evidenced in the design and could not be exploited for model testing. However, the limitation was not crucial, as evidenced by the joint frequency distribution for the properties $A; P' : U$ and $A; U : P'$ shown by Table 7. First, there is a greater number of subjects satisfying both independence properties $A; P' : U$ and $A; U : P'$ than random behavior would predict. This clustering of subjects was also noted in Table 6. However, Table 7 implies a secondary clustering of subjects whose

TABLE 7

Observed Frequencies and Expected Frequencies under Random Model of Values $\bar{r}_s$ for Properties $A; U : P'$ and $A; P : U$

	$A; U : P'$			
$A; P' : U$	$[-1, -.83]^a$	$[-.67, -.17]$	$[0, .67]$	$[.83, 1]^b$
		Actual Frequencies		
$[-.50, -.17]^a$	0	0	3	7
$[-.08, .67]$	0	1	8	3
$[.83, 1]^b$	0	0	0	9
		Expected Frequencies		
$[-.50, -.17]^a$	.50	5.23	6.13	.50
$[-.08, .67]$	.46	7.44	9.98	.46
$[.83, 1]^b$	.05	.10	.10	.05

[a] Values consistent with sign dependence version of this property within one pairwise inversion.
[b] Values consistent with independence version of this property within one pairwise inversion.

reponses are now more interpretable. These are subjects who satisfied independence $A; U : P'$ and satisfied sign dependence $A; P^* : U$. This is evidenced by the seven subjects (vs. .50 expected if subjects responded randomly) with values of $\bar{r}_s \geqslant .83$ for the property $A; U : P'$ and values of $\bar{r}_s \leqslant -.17$ for the property $A; P' : U$.

An important implication of this secondary clustering of subjects is that it represents a group of subjects who showed a sign dependence $A; P^* : U$. This behavior is inconsistent with all the three-factor polynomial models of choice under ambiguity in Table 1. None of the three-factor models incorporates a term P which is signed, since P is certainly not signed in its multiplicative relationship with U. If it were, then, over a range of values for the probability P, a subject would prefer a lower probability of receiving a particular gain having positive utility. This is completely implausible, and therefore all of the three-factor models must be rejected.

The property $A; P^* : U$ for these subjects in their choice behavior under ambiguity implies that they were reversing their preferences over ambiguity, depending upon the probability level. This process could underlie the observation by Curley and Yates (1985) of a positive correlation between the probability level P and the extent of ambiguity avoidance by their subjects as a group, as well as similar findings by other researchers (Einhorn & Hogarth, 1986; Goldsmith & Sahlin, 1982; Hogarth & Kunreuther, 1985). Moreover, the signedness of P as a characteristic of subjects' behavior may be a more general phenomenon. Yates and Carlson (1986) hypothesized and found evidence for a similar signedness of likelihood judgments within a task of ordering future real-world events by their judged likelihood. In addition, computerized expert systems, for example, MYCIN (Shortliffe, 1976) and CASNET (Weiss, Kulikowski, Amarel, & Safir, 1978), have used signed likelihood functions in an attempt to better approximate experts' reasoning. The signedness of P, in its relation with preference over ambiguity A, is therefore a major finding of Study 2.

Portfolio Theory. In addition to Models 3–6, Study 2 tested a variant of portfolio theory as applied to choice under ambiguity. Portfolio theory can be adapted to describe choice under ambiguity by including ambiguity as a factor which influences the undefined parameter of perceived risk. A principal feature of the theory is that preference is assumed to be single-peaked over risk when the level of expected value is fixed. Extending this viewpoint, preference would be similarly single-peaked over ambiguity when the expected value is fixed. This single-peakedness could result from an approach–avoidance conflict, for example, between the good of an increasing P_{max} and the bad of a decreasing P_{min}, as ambiguity increases (Coombs & Avrunin, 1977). Another prediction of the extended theory is that a subject should be indifferent between options sharing both the same expectation and the same level of risk/ambiguity.

These predictions were tested in Study 2. Recall that a lottery having $x = \$5$, $P = .50$, and $R = .50$ was included in the design. This lottery has the same expectation and degree of ambiguity as the lottery having $x = \$10$, $P = .25$, and $R = .50$, another of the lotteries used. Assuming that their degree of risk was the same in that no losses were involved, indifference should have obtained between these lotteries. For each subject, there were five replications of this lottery pair. Of 31 subjects, 13 (41.9%) showed the same preference over all replications, and 10 (32.3%) had four of the five agree. This distribution of preferences significantly differed from the binomial distribution, expected with random responses under indifference ($\chi^2 = 101.75$, $p < .0001$), that was predicted by the present variation of portfolio theory.

Also telling with respect to the theory is the prediction of single-peakedness over ambiguity. If this aspect of the theory is useful, the number of single-peaked preference patterns would be expected to significantly exceed zero. For each subject in Study 2, there were six preference patterns over levels of the factor A, one at each joint level of the factors U and P, for which expectation was fixed. Each pattern could be classified as "monotonic," if preference either increased or decreased with the level of ambiguity, "single-peaked," if the lottery having $R = .20$ was most preferred, or "single-dipped," if the lottery having $R = .20$ was least preferred. The latter pattern is inconsistent with portfolio theory, and so can be used in comparison with the single-peaked pattern. Specifically, according to the theory, the frequency of single-peaked patterns should have exceeded the frequency of single-dipped patterns. Of 186 possible patterns, 48 were not monotonic. Of these 48, 25 (52.1%) were single-peaked and 23 were single-dipped. These values did not significantly differ (binomial, $z = .14$, $p > .8$).

Thus, neither test of the relevant data from Study 2 lends support to portfolio theory as operative under ambiguity. The possibility of portfolio theory being descriptive of choice under ambiguity does not appear, therefore, to offer promise.

General Discussion

Summary

Ambiguous options were defined as those having uncertain outcome-generating processes, and were characterized as having uncertainty about the probabilities associated with the possible outcomes. Choices involving such options have manifested a significant degree of reaction to their ambiguity, typically ambiguity avoidance. The present research primarily addressed the descriptive modeling of individuals' ambiguity reactions in choice situations.

A descriptive model, if it can be constructed, may or may not correspond to any normative or prescriptive models that could be adopted to guide one's choices. For example, the model of statistical decision theory has been espoused as a normative model of choice, even though it clearly does not describe ambiguity reactions (Raiffa, 1961; Roberts, 1963; Savage, 1972). Still, in that rationality is a criterion of normativeness, capturing ambiguity reactions by a descriptive model may aid the evaluation of the normative value of ambiguity reactions (Fellner, 1961). In addition, a descriptive model can illuminate the processes whereby subjects evaluate alternatives in situations involving ambiguity. In so doing, the model would have value for the design of decision aids and could provide insight into the structure of realistic choice situations.

In pursuit of the goal of description, a number of plausible models for choice under ambiguity were isolated. These models were general forms of models that have been proposed in the literature, and are listed as Models 2–6 in Table 1. Two studies were performed for the purpose of testing some of the necessary implications of these models, so as to differentiate among them. Study 1 addressed Model 2, the lexicographic model, and derived estimates of the extent of ambiguity reactions in a choice situation. Study 2 addressed Models 3–6, the three-factor polynomial models, and a model adapted from portfolio theory.

Study 1. The lexicographic model of choice under ambiguity, in which expectation serves as the first dimension and ambiguity the second, was not supported by Study 1. Significant reactions to ambiguity were obtained in the neighborhood of all three levels of the best-guess probability P used, $P=.25$, $P=.50$, and $P=.75$. This reaction was predominantly that of ambiguity avoidance at the higher levels of P, but significant ambiguity seeking at $P=.25$ was identified. The finding of ambiguity-seeking behavior was an important confirmation of a conjecture that such behavior exists.

Subjects in Study 1 were willing to trade off a substantial percentage, on the order of 5–10%, of their expected gain in their reactions to ambiguity. At $P=.50$, the magnitude of the tradeoff elicited from subjects' choice responses, was comparable to that which has been obtained using pricing procedures.

Study 2. Four generalized three-factor models of choice under ambiguity were tested. Diagnostic necessary properties of these models were identified in the context of conjoint measurement theory and capitalized upon in the design of Study 2.

Subjects exhibited a consistency of behavior in Study 2 which allowed for the analyses. However, the manipulation of the factor U was not successful. Subjects did not behave as if \$0 was a reference point against which the utilities of other outcomes were measured; their responses were not consistent with a negative value being attached to negative amounts. This conclusion apparently reduced the study's ability to differentiate the models, except that a significant sign dependence $A; P^* : U$ obviated the need for the manipulation. The identification of subjects exhibiting the sign dependence $A; P^* : U$ was the major finding of Study 2. None of the three-factor models can accommodate the observed sign dependence. Thus, they, like the lexicographic model, were eliminated as plausible general models of choice under ambiguity.

Similarly, the model describing choice as an adaptation of portfolio theory was not supported. Subjects were not indifferent between lotteries having equal expectation and ambiguity. And subjects did not evidence a significant degree of single-peaked preference over ambiguity.

Modeling Implications

Since the majority of models that have been proposed in the literature to describe ambiguity reactions are special cases of the tested models, the research has efficiently excluded these more specific versions as possible descriptive theories, as well. This elimination has furthered the development of our knowledge regarding the goal of descriptively modeling subjects' choice behavior under ambiguity, albeit in a somewhat negative fashion. In light of the results, what are the prospects for developing a descriptive choice model?

Several possibilities still exist for the construction of a satisfactory model. First, a model using four or more factors may be both workable and descriptive. For example, models proposed by Einhorn and Hogarth (1985; 1986) have been discussed which use more than three factors within algebraic combination rules. Other models having more than three factors, besides these, can also be constructed which both generalize the statistical decision theory model and accommodate the property $A; P^* : U$. An investigation of the models in the spirit of Study 2 might illuminate their applicability.

A second possibility is to modify the function denoted by the factor A. Perhaps A should be conceived as a complex function which depends upon more than just ambiguity; for example, as being influenced by the probability level and the magnitude and type of outcomes involved. Whereas the four-factor models take the approach of complicating the composition rule, this approach complicates a function within the simpler three-factor composition rules. Both approaches effectively involve adding terms to the polynomial model forms.

A third possibility is to abandon the attempt of capturing decision makers' behavior by a single algebraic form. A repertoire of model strategies used by different individuals in different settings might describe decisions better. This approach complicates the proposed metastrategy of subjects. Its lack of parsimony makes this a less attractive possibility, a priori.

Finally, we might attempt to rethink the model, qualitatively, and build a new model which is not generalized from statistical decision theory. Perhaps the use of the lottery as a representation of realistic decision situations is itself inadequate, and a more appropriate representation will lead to the construction of a more appropriate model form.

Aside from implying the inadequacy of a three-factor model of choice, another significant implication of the observed sign dependence $A; P^* : U$ in Study 2 concerns the likelihood function. One hypothesis which is consistent with the sign dependence is that subjects used a likelihood function which could assume negative as well as positive values. It was as if events were categorized as being likely or unlikely, with unlikely events treated as having negative likelihoods. This is in contrast to the probability function as a measure of likelihood, which has a strictly nonnegative range. The implications of a likelihood function of this form are considerable, in terms of both its impact on the psychology of descriptive choice theory and its impact on practical decision making and decision aids.

Acknowledgments

These experiments were undertaken by the first author as part of his dissertation at the University of Michigan in partial fulfillment of the requirements for the Ph.D. in Psychology. Also, an earlier version of this paper was presented at the 19th Annual meeting of the Society for Mathematical Psychology. Cambridge, MA, August 20, 1986. The authors are thankful for the comments on the earlier drafts, particularly from the members of the doctoral committee: Richard G. Cornell, David E. Meyer, J. E. Keith Smith, and Mark J. Young. We also thank Michele Watrin for her help in preparing the manuscript. The research was supported in part by National Institute of Mental Health Training Grant MH16892.

References

Becker, S. W., & Brownson, F. O. (1964). What price ambiguity? Or the role of ambiguity in decision-making. *Journal of Political Economy*, **72**, 62–73.

Budescu, D. V., & Wallsten, T. S. (1985). Consistency in interpretation of probabilistic phrases. *Organizational Behavior and Human Decision Processes*, **36**, 391–405.

Cochran, W. G., & Cox, G. M. (1957). *Experimental designs* (2nd ed.). New York: Wiley.

Coombs, C. H., & Avrunin, G. S. (1977). Single-peaked functions and the theory of preference. *Psychological Review*, **84**, 216–230.

Coombs, C. H., & Bowen, J. (1971). Additivity of risk in portfolios. *Perception and Performance*, **10**, 43–46.

Coombs, C. H., Donnell, M. L., & Kirk, D. B. (1978). An experimental study of risk preference in lotteries. *Journal of Experimental Psychology: Human Perception and Performance*, **4**, 497–512.

Coombs, C. H., & Huang, L. C. (1970). Tests of a portfolio theory of risk preference. *Journal of Experimental Psychology*, **85**, 23–29.

Coombs, C. H., & Meyer, D. E. (1969). Risk-preference in coin-toss games. *Journal of Mathematical Psychology*, **6**, 514–527.

Curley, S. P. (1986). Descriptive models of choice under ambiguity [Doctoral dissertation, University of Michigan, 1986]. *Dissertation Abstracts International*, **47**, 1323B.

CURLEY, S. P. (1989). *Practical application of axiomatic conjoint measurement.* Unpublished manuscript, University of Minnesota, Minneapolis.

CURLEY, S. P., ERAKER, S. A., & YATES, J. F. (1984). An investigation of patients' reactions to therapeutic uncertainty. *Medical Decision Making*, **4**, 501-511.

CURLEY, S. P., & YATES, J. F. (1985). The center and range of the probability interval as factors affecting ambiguity preferences. *Organizational Behavior and Human Decision Processes*, **36**, 273–287.

CURLEY, S. P., YOUNG, M. J., & YATES, J. F. (1989). Characterizing physicians' perceptions of ambiguity. *Medical Decision Making* **9**, 116–124.

EINHORN, H. J., & HOGARTH, R. M. (1985). Ambiguity and uncertainty in probabilistic inference. *Psychological Review*, **92**, 433–461.

EINHORN, H. J., & HOGARTH, R. M. (1986). Decision making under ambiguity. *Journal of Business*, **59**, S225–S250.

EINHORN, H. J., & HOGARTH, R. M. (1988). Decision making under ambiguity: A note. In B. R. Munier (Ed.), *Risk. decision, and rationality.* Dordrecht, Holland: Reidel.

ELLSBERG, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, **75**, 643–669.

ELLSBERG, D. (1963). Risk, ambiguity, and the Savage axioms: Reply. *Quarterly Journal of Economics*, 77, 336–342.

FELLNER, W. (1961). Distortion of subjective probabilities as a reaction to uncertainty. *Quarterly Journal of Economics*, **75**, 670–689.

FISHBURN, P. C. (1980). A simple model for the utility of gambling. *Psychometrika*, **45**, 435–448.

GARDENFORS, P. (1979). Forecasts, decisions and uncertain probabilities. *Erkenntnis*, **14**, 159–181.

GARDENFORS, P., & SAHLIN, N. E. (1982a). Reply to Levi. *Synthese*, **53**, 433–438.

GARDENFORS, P., & SAHLIN, N. E. (1982b). Unreliable probabilities, risk taking, and decision making. *Synthese*, **53**, 361–386.

GARDENFORS, P., & SAHLIN, N. E. (1983). Decision making with unreliable probabilities. *British Journal of Mathematical and Statistical Psychology*, **36**, 240–251.

GOLDSMITH, R. W., & SAHLIN, N. E. (1982). The role of second-order probabilities in decision making. In P. Humphreys, O. Svenson, & A. Vari (Eds.), *Analysing and aiding decision processes.* New York: North-Holland.

GOLDSTEIN, W. M. (1985). The relationship between judgment and choice [Doctoral dissertation, University of Michigan, 1984]. *Dissertation Abstracts International*, **45**, 2341B.

GOLDSTEIN, W. M., & EINHORN, H. J. (1987). Expression theory and the preference reversal phenomena. *Psychological Review*, **94**, 236–254.

GRETHER, D. M., & PLOTT, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *American Economic Review*, **69**, 623–638.

HOGARTH, R. M., & KUNREUTHER, H. (1985). Ambiguity and insurance decisions. *American Economic Review (AEA Papers and Proceedings)*, **75**, 386–390.

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

KAHNEMAN, D., & TVERSKY, A. (1973). On the psychology of prediction. *Psychological Review*, **80**, 237–251.

KAHNEMAN, D., & TVERSKY, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, **47**, 263-291.

KENDALL, M. G. (1970). *Rank correlation methods* (4th ed.). London: Griffin.

KEYNES, J. M. (1921). *A treatise on probability.* London: MacMillan.

KNIGHT, F. H. (1921). *Risk, uncertainty and profit.* Boston: Houghton Mifflin.

KRANTZ, D. H., LUCE, R. D., SUPPES, P., & TVERSKY, A. (1971). *Foundations of measurement* (Vol. 1). New York: Academic Press.

KRANTZ, D. H., & TVERSKY, A. (1971). Conjoint-measurement analysis of composition rules in psychology. *Psychological Review*, **78**, 151–169.

LEE, W. (1971). *Decision theory and human behavior.* New York: Wiley.

LIBBY, R., & FISHBURN, P. C. (1977). Behavioral models of risk taking in business decisions: A survey and evaluation. *Journal of Accounting Research*, **15**, 272–292.

LICHTENSTEIN, S., & SLOVIC, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, **89**, 46–55.

LOPES, L. L. (1983). Some thoughts on the psychological concept of risk. *Journal of Experimental Psychology: Human Perception and Performance*, **9**, 137–144.

LUCE, R. D. (1980). Several possible measures of risk. *Theory and Decision*, **12**, 217–228.

LUCE, R. D., & TUKEY, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, **1**, 1–27.

POLLATSEK, A., & TVERSKY, A. (1970). A theory of risk. *Journal of Mathematical Psychology*, **7**, 540–553.

RAIFFA, H. (1961). Risk, ambiguity, and the Savage axioms: Comment. *Quarterly Journal of Economics*, **75**, 690–694.

ROBERTS, H. V. (1963). Risk, ambiguity, and the Savage axioms: Comment. *Quarterly Journal of Economics*, **77**, 327–336.

SAVAGE, L. J. (1972). *The foundations of statistics* (2nd rev. ed.). New York: Dover.

SCHOEMAKER, P. J. H. (1982). The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature*, **20**, 529–563.

SHAFER, G. (1976). *A mathematical theory of evidence*. Princeton, NJ: Princeton Univ. Press.

SHAFER, G., & TVERSKY, A. (1985). Languages and designs for probability judgment. *Cognitive Science*, **9**, 309–339.

SHORTLIFFE, E. H. (1976). *Computer-based medical consultations: MYCIN*. New York, Elsevier.

SLOVIC, P. (1964). Assessment of risk-taking behavior. *Psychological Bulletin*, **61**, 220–233.

SLOVIC, P. (1966). Value as a determiner of subjective probability. *IEEE Transactions on Human Factors in Electronics*, **7**, 22–28.

SMITH, V. L. (1969). Measuring nonmonetary utilities in uncertain choices: The Ellsberg urn. *Quarterly Journal of Economics*, **83**, 324–329.

TODA, M., & SHUFORD, E. H., Jr. (1965). Utility, induced utilities, and small worlds. *Behavioral Science*, **10**, 238–254.

TVERSKY, A., & KAHNEMAN, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, **76**, 105–110.

WEBER, E. U. (1984). Combine and conquer: A joint application of conjoint and functional approaches to the problem of risk measurement. *Journal of Experimental Psychology: Human Perception and Performance*, **10**, 179–194.

WEINSTEIN, M. C., FINEBERG, H. V., ELSTEIN, A. S., FRAZIER, H. S., NEUHAUSER, D., NEUTRA, R. R., & MCNEIL, B. J. (1980). *Clinical decision analysis*. Philadelphia: Saunders.

WEISS, S. M., KULIKOWSKI, C. A., AMAREL, S., & SAFIR, A. (1978). A model-based method for computer-aided medical decision making. *Artificial Intelligence*, **11**, 145–172.

WINKLER, R. L. (1972). *Introduction to Bayesian inference and decision*. New York: Holt, Rinehart & Winston.

YATES, J. F., & CARLSON, B. W. (1986). Conjunction errors: Evidence for multiple judgment procedures, including "signed summation." *Organizational Behavior and Human Decision Processes*, **37**, 230–253.

YATES, J. F., & ZUKOWSKI, L. G. (1976). Characterization of ambiguity in decision making. *Behavioral Science*, **21**, 19–25.

ZADEH, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, **1**, 3–28.

RECEIVED: September 25, 1987