

Scale Convergence and Utility Measurement

MICHAEL H. BIRNBAUM AND SARA E. SUTTON

California State University, Fullerton

In order to investigate derived scales for the utility, or subjective value of money, subjects were instructed to perform four tasks: in two tasks, they judged "ratios" and "differences" of strengths of preference for monetary amounts; in two other tasks, they judged the values of gambles from buyer's or seller's points of view. The two arrays of data for "ratios" and "differences" were consistent with the hypothesis that most subjects used only one operation to compare monetary amounts, although a few subjects appeared to use two operations. The buyer's and seller's prices would lead to two different utility functions for money under expected utility theory, subjective expected utility theory, or any theory that is additive across outcomes. However, configural-weight utility theory can predict these changes in rank order with an invariant utility function, by postulating that the configural weight of the smaller amount depends on point of view. The data also reveal systematic violations of dominance (monotonicity) that can be described by assuming that the configural weight of zero, when it is the lower value, has smaller weight at low probabilities than nonzero outcomes. Disregarding the minority of subjects who appeared to utilize two operations for judging "ratios" and "differences" in utility, the majority of the data can be well-approximated using a single scale of utility, the subtractive model for "ratio" and "difference" comparisons, and the configural weight, rank-dependent model for buyer's and seller's prices. © 1992 Academic Press, Inc.

There is a great deal of evidence that expected utility theory (Stigler, 1950a,b; von Neumann & Morgenstern, 1947) cannot account for the observed preference orderings of gambles (Allais, 1979; Birnbaum, Coffey, Mellers & Weiss, 1992; Edwards, 1954; Ellsberg, 1961; Fishburn, 1970, 1983; Kahneman & Tversky, 1979; Keeney & Raiffa, 1976; Machina, 1982; Payne, 1973; Schoemaker, 1982; Slovic, Lichtenstein, & Fischhoff, 1988; von Winterfeldt & Edwards, 1986). Some of the deviations between what expected utility theory dictates that decision makers should do and what humans actually do when confronted with choices can be explained by subjective expected utility theory (Savage, 1954), prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1986),

This research was supported by a Hughes Fellowship and a Senior Faculty Fellowship from California State University, Fullerton. We thank R. Duncan Luce, Colleen Moore, and Mary Kay Stevenson-Busemeyer for their comments on an earlier draft. Requests for reprints should be addressed to Michael H. Birnbaum, Department of Psychology, California State University, Fullerton, CA 92634. Sara Sutton is now at the University of Rochester.

lottery-dependent utility theory (Becker & Sarin, 1987), weighted utility theory (Karmarkar, 1978), and other generalizations (Miyamoto, 1989).

Luce and Narens (1985) developed dual-bilinear utility theory as the most general representation of its class that would lead to an interval scale of utility and showed that it can account for evidence that has been taken as refutation of expected utility theory (see also Narens & Luce, 1986). Luce (1986) noted that dual-bilinear utility theory has not yet been disproved by the phenomena that have posed difficulty for the subjective expected utility theories. This theory is closely related to rank-dependent utility theory (Quiggin, 1982; Karni & Safra, 1987; Luce & Fishburn, 1991; Yaari, 1987) and also to configural weight theory (Birnbaum, 1974a; Birnbaum, Parducci, & Gifford, 1971; Birnbaum & Stegner, 1979, 1981; Birnbaum & Veit, 1974).

Configural weight theory allows the weight of a stimulus to depend on its position among the configuration of stimuli to be considered. This theory was developed to account for deviations from additive and parallel averaging models in psychophysical and social evaluative judgment tasks, while preserving the agreement, or convergence, of scales defined by models applied to several phenomena (Birnbaum *et al.*, 1971; Birnbaum, 1974a; Birnbaum & Veit, 1974; Birnbaum & Stegner, 1979, 1981). Although the general concept of configural weighting allows many complex patterns, these experiments were generally consistent with the simplest form, the range model. As Birnbaum (1974a) and Birnbaum and Stegner (1979) noted, the range model can be interpreted as a rank-dependent weighted averaging model. When this model was fit to the data, it yields scales of subjective value that agree with scales derived from "difference" ratings (Birnbaum, 1974a, 1982).

The principle of scale convergence states that when considering rival theories proposed to describe different empirical phenomena involving the same theoretical constructs, preference should be given to coherent theoretical systems (in which the same measurement scales can be used to account for a variety of empirical phenomena) as opposed to theoretical systems that require different measurements for each new situation.

In the case of utility theory, it has been found that expected utility theory yields inconsistent scales of utility and that these scales seem to differ systematically from scales obtained using other methods for scaling subjective value (Birnbaum *et al.*, 1992; Bell & Raiffa, 1988; Becker & Sarin, 1987; Edwards, von Winterfeldt, & Moody, 1988; von Winterfeldt & Edwards, 1986; Keller, 1985; Hershey, Kunreuther, & Schoemaker, 1982). There has been theoretical disagreement among these authors whether to postulate that utility and value are different scales, or whether to revise theories in order to achieve unity of scales.

Birnbaum (1987) noted that the deviations from scale convergence that

occur for utility are similar in pattern to the deviations that occur by forcing the assumption of additivity (or parallel averaging) in psychophysical integration and evaluative judgment. One can achieve scale convergence in information integration by giving up the assumption of additive or parallel-averaging models in favor of configural weighting (Birnbbaum, 1974a, 1982; Birnbbaum & Jou, 1990; Birnbbaum, *et al.*, 1971; Birnbbaum & Veit, 1974). Configural weighting theory has the hope of resolving the inconsistent scales for utility and value measurement by separating the scaling of stimuli from the scaling of uncertainty and risk.

Birnbbaum and Stegner (1979) found that manipulation of the judge's point of view from buyer to seller produced a change in preference order that could be explained by the assumption that the manipulation affected the configural weight. Buyer's and seller's judgments of value of gambles have been employed by Coombs, Bezeminder, and Goode (1967) and by Lichtenstein and Slovic (1971), to check the generality of other phenomena.

Buyer's and seller's prices are not expected to be equal, except under special circumstances (Raiffa, 1968), because buyers and sellers can reasonably be theorized to be working from different levels of wealth on the utility function. If the utility function is negatively accelerated, as is usually assumed, then changes in wealth will make less of a difference to a person at a higher level of wealth than to persons of lower wealth. Giving a lottery to a subject increases the subject's wealth; hence selling prices theoretically reflect the utilities of a wealthier person than buyer's prices.

However, the difference observed between buyer's and seller's prices is too large to be explained by such an "income effect," as noted for example by Knetsch and Sinden (1984), who ascertained willingness to buy or sell lottery tickets when either money or a ticket was given to the subjects. Since their subjects were randomly assigned to either receive the money or the ticket, the proportion of subjects who preferred the money or ticket should have been independent of which one had been given initially. Instead, subjects were reluctant to exchange, yielding higher selling than buying prices.

Birnbbaum *et al.* (1992) extended the analysis of the buyer-seller distinction in Birnbbaum and Stegner (1979) and found that manipulating the buyer's and seller's points of view also changed the preference order for gambles as if the configural weighting for the higher or lower valued outcomes was changed, and the utility function was invariant.

Birnbbaum *et al.* (1992) observed phenomena that posed difficulties for additive models. Consider the following generalization of subjective expected utility theory:

$$G = M[S_1(p_1)u_1(x) + S_2(p_2)u_2(y)], \quad (1)$$

where G is the overall value of the gamble to receive outcomes x or y with probabilities p_1 and p_2 , respectively; M is a strictly increasing monotonic function; $S_1(p_1)$ and $S_2(p_2)$ are the decision weights that depend on the probabilities p_1 and p_2 ; and $u_1(x)$ and $u_2(y)$ are subjective values of the monetary outcomes, x and y , respectively.

The rank order of gambles observed by Birnbaum *et al.* (1992) changed in such a fashion that Eq. (1) would require two different scales of u for buyer's and seller's points of view. However, Birnbaum *et al.* (1992) found that it was possible to describe the changes in rank order using a single scale of u , by means of configural weight theory, in which only the configural weight parameter depended on point of view.

Dual-bilinear theory, the simplest form of configural weighting, allows Eq. (1) to be broken into two related parts, for the cases where $x > y$ or $x < y$. If Eq. (1) applies when $x < y$, then the functions S_1 and S_2 trade places for the case when $x > y$. Although this theory seems a minor generalization of the additive model, it yields very different utility functions when applied to the same data (Birnbaum *et al.*, 1992), and it allows utility scales estimated separately from the case where $x < y$, and from the case where $x > y$ to be different from those estimated when x and y are both free to vary across the domain. For further information on the relationship between scales derived with and without the configural weighting assumption, see Birnbaum (1981, 1982).

The purpose of the present research is to further test the implications of the utility functions implied by configural weighting theory and to compare different methods for scaling the utility or subjective value of money. In these experiments, subjects judge buyer's and seller's prices of gambles, as in Birnbaum *et al.* (1992), using a design chosen to facilitate derivation of a well-defined utility function. In addition, subjects judge "ratios" and "differences" of preferences for receiving different monetary amounts, to permit estimation of riskless scales of strength of preference.¹

THEORY

Point of View

For gambles that have subjectively equally likely 50–50 outcomes, the configural weight model can be written as follows:

$$G_{\sqrt{}}(x.5y) = u^{-1} [.5(u(x) + u(y)) + W_{\sqrt{}}|u(x) - u(y)|], \quad (2)$$

¹ Quotation marks are used to distinguish instructions to judge "ratios" and "differences" and the judgments obtained under such instructions from actual or theoretical ratios and differences. This distinction is necessary because judged "ratios" and "differences" may or may not fit ratio and difference models.

where $G_V(x, .5y)$ is the predicted judgment in point of view V for a gamble to win x with probability $.5$ and otherwise to receive y ; u is the utility function, and W_V is the configural weight. The configural weight can be thought of as the amount of weight taken from the lower-valued stimulus and given to the higher-valued stimulus (when W_V is positive) or taken from the higher valued and given to the lower valued stimulus (when W_V is negative) within each gamble. When W_V is $-.5$ or $.5$, all of the weight is given either the lower or higher stimulus, respectively.

Each panel of Fig. 1 illustrates predicted values of 50-50 gambles between two amounts, x and y , according to configural-weight theory. Each panel shows predictions for a different value of the configural weight parameter, using $u(x) = x^{.70}$ for the utility function. The predictions in Fig. 1 are based on the simplifying assumptions that the response function (transforming utility into a monetary judgment) is the inverse of the utility function for money and that the subjective weight of each outcome in these 50-50 gambles is $.5$.

The predictions are illustrated for the case, $u(x) = x^{.70}$. The upper, middle, and lower panels show predictions for $W_V = -.25, 0$, and $.25$, respectively. Within each panel, predicted values are plotted as a function of Amount x , with a separate curve for each level of Amount y . When $W_V = 0$, there is a slight divergence, due to the fact that $u(x)$ is concave downwards (and therefore, its inverse function converts the otherwise parallel curves to divergent ones). The rank order of the points depends on both W_V and the utility function. If $u(x)$ is linear, then W_V can cause divergence or convergence (risk avoidance or risk seeking), depending on whether W_V is less than or greater than zero. In extreme cases, when $W_V = -.5$ or $.5$, Eq. (2) becomes a minimum or maximum model, respectively.

Consideration of Fig. 1 and Eq. 2 thus reveals that "risk aversion" and "risk seeking" in this model are not equivalent to the shape of the utility function, but can be described in part by the configural weighting, and in part by the utility function. The next section shows that what has been called "risk aversion" or "risk seeking" can also be interpreted as a consequence of the function relating weighting to subjective probability.

Dominance Violations

Birnbaum *et al.* (1992) found systematic violations of dominance (monotonicity). Let $G_V(xpy)$ represent the value of a gamble to win x with probability p or win y with probability $1 - p$, from point of view, V . Monotonicity requires that if x or y or both were increased, the value of the gamble should increase. Instead, Birnbaum *et al.* (1992) found that from all three points of view, $G_V(\$96, .95, \$0)$ was higher on the average than $G_V(\$96, .95, \$24)$, even though the latter gamble dominates the

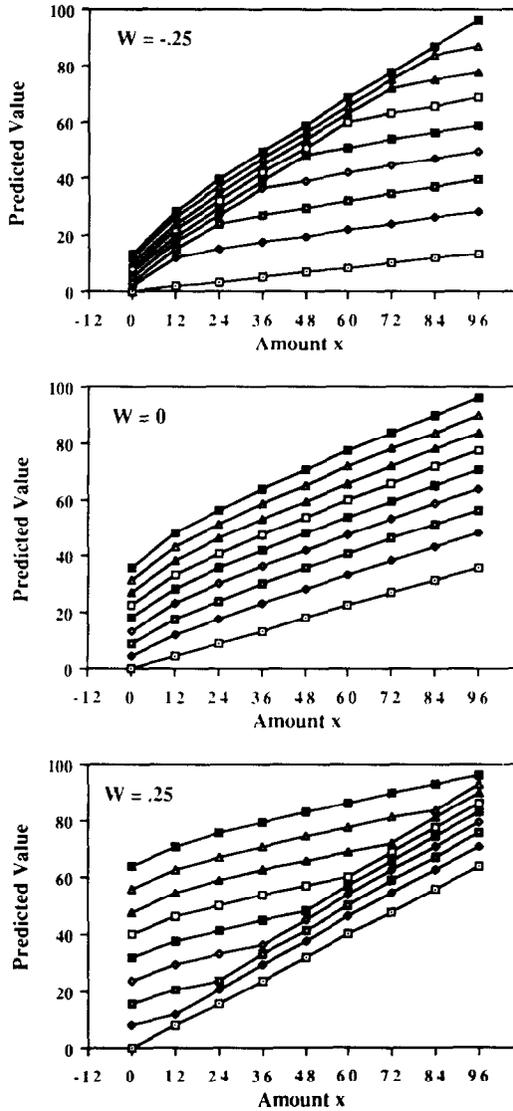


FIG. 1. Predicted values according to configural weighting theory [Eq. (2)], illustrated using $u(x) = x^{.70}$ and $W_v = -.25, 0, \text{ and } .25$. In each panel, values are plotted as a function of Amount x with a separate curve for each level of Amount y .

former. As Luce (1986) noted, the dual-bilinear representation [as in Eq. (2)] does not allow such violations of monotonicity, nor does the rank and sign dependent theory of Luce and Fishburn (1991). However, a configural weighting theory was developed to describe the result; Birnbaum *et*

al. (1992) explained this effect by theorizing that the outcome \$0 is given a lower weight than non-zero outcomes, especially when the probability of \$0 is small.

To illustrate the configural weighting model of Birnbaum *et al.* (1992), let $S_0(p)$ and $S_L(p)$ represent the configural weighting functions of probability for the lower-valued outcome of \$0 and for the lower-valued outcome, x , where $0 < x < y$. For these gambles, this extension of Eq. (2) can be written:

$$G_V(0py) = u^{-1}[S_0(p)u(0) + (1 - S_0(p))u(y)]$$

$$G_V(xpy) = u^{-1}[S_L(p)u(x) + (1 - S_L(p))u(y)].$$

To illustrate how this theory can predict violations of dominance, Fig. 2 plots predictions for a numerical example that also uses $u(x) = x^{.70}$. In this illustration, the weighting functions are $S_0(p) = .22 + .65p$ and $S_L(p) = .50 + .38p$, for $.04 < p < .96$. The crossing of curves in Fig. 2 indicates violation of dominance.

The pattern in Fig. 2 was obtained empirically for all three points of view in the study by Birnbaum *et al.* (1992), and it has since been replicated with different subjects, different stimulus formats, with different stimulus levels (including losses), and even when subjects have financial incentives (Mellers, Weiss, & Birnbaum, 1992). It seem unlikely, however, that subjects would prefer the dominated gamble when given a direct choice. To check this possibility, choice data were also obtained in the present study.

These configural weighting functions for the probability of the lower-valued outcome also show how low probability gambles to win large prizes (e.g., like typical state lottery tickets) might be valued above expected value, even when the u function is concave downwards and even

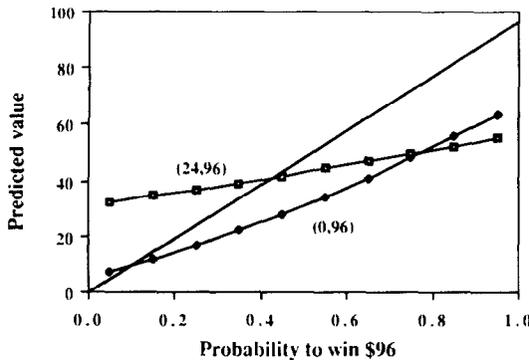


FIG. 2. Predicted violation of dominance, based on differential weighting of the Amount $x = \$0$.

when the subject places greater weight on the lower-valued outcomes. The line in Fig. 2 with the steepest slope represents the expected values of the (\$96, p , \$0) gambles. Note that the predicted judgment of (\$96, .05, \$0) exceeds its expected value. In this model, the configural weights can produce what appears to be "risk seeking," even when the utility function is negatively accelerated and even when the lower valued outcomes receive greater weight than higher valued outcomes, if the weight of small probabilities is sufficiently large compared to objective probability.

"Ratios" and "Differences"

The utility function defined by the configural-weight theory is theorized to be independent of the judge's point of view (Birnbbaum *et al.*, 1992). By the principle of scale convergence, the theory would be further enhanced if its u -function were able to predict other phenomena, such as judgments of strength of preference.² For the same monetary amounts, x and y , we can ask subjects to rate the "difference" in utility and the "ratio" of utilities between the riskless outcomes.

Scaling the utility of money "directly" has been attempted in several investigations (Stevens, 1975; Galanter, 1962; Breault, 1983). In the "direct" scaling approach, subjects are asked to assign numbers to represent their utilities "directly." However, "direct" scaling has the disadvantage that the numbers assigned by subjects constitute in principle only an ordinal scale and place few constraints on theories of decision making, unless the numbers can be taken at "face value" (Birnbbaum, 1978, 1980, 1982). There is a great deal of evidence that "direct" estimations of "ratios," for example, cannot be taken at "face value" (Birnbbaum, 1982; Hardin & Birnbbaum, 1990; Mellers & Birnbbaum, 1982, 1983).

In the present study, however, the experiments are designed so that "direct" judgments comparing riskless utilities can be investigated within the structure of testable models that permit estimation of scales from ordinal constraints in the data in order to define scales that are unique to at least an interval scale (Birnbbaum, 1978, 1980, 1982), allowing a non-trivial test of scale convergence between the two types of judgments.

Suppose judgments of "ratios" and "differences" conform to the following models:

² To some theorists, the term "utility" applies only to risky situations; therefore, riskless utility would be a contradiction in terms. These authors distinguish the concepts of utility and value, and they often use the notations u and v to represent the two postulated scales (e.g., Bell & Raiffa, 1988). In the present approach, the intention is to test theories that satisfy scale convergence, the premise that the same measures mediate different empirical relations; therefore, the notation u is used throughout to emphasize the assumption of scale convergence.

$$R_{xy} = J_R[u(x)/u(y)] \quad (3)$$

$$D_{xy} = J_D[u(x) - u(y)], \quad (4)$$

where R_{xy} and D_{xy} are judged "ratios" and "differences," respectively; J_R and J_D are strictly monotonic judgment functions that convert subjective ratios of utility [$u(x)/u(y)$] and differences in utility [$u(x) - u(y)$] into overt judgments; and $u(x)$ is the utility function. If subjects judge both operations on the same scale, then the two judgments would not be monotonically related in general, but instead would show a particular interrelationship (Birnbau, 1980). If this theory is compatible with the data, then in principle, scales of utility or subjective value could be determined to a ratio scale (Krantz, Luce, Suppes, & Tversky, 1971; Miyamoto, 1983).

Figure 3 illustrates predictions for the theory that subjects use both operations on the same scale of utility [Eqs. (3 and 4)], using $u(x) = x^{.70}$ for the illustration. The upper panel plots predicted "ratios" [ratios of utility multiplied by 100; $J_R(r) = 100r$] as a function of Amount x , with a separate curve for each level of Amount y . The middle panel plots predicted "differences" [differences of utility multiplied by .801, to put them on a scale from -15 to 15; i.e., $J_D(d) = .801d$] as a function of Amount x with a separate curve for each level of Amount y . The lower panel plots predicted "ratios" against predicted "differences." Note that the rank order of predictions in Fig. 3 differs systematically. For a given difference, ratios approach 1 as the difference is moved up the scale; for example, $2 - 1 = 3 - 2 = 4 - 3$, but $2/1 > 3/2 > 4/3$. In Fig. 3, note that predicted "ratios" have the same property: For a given difference, the open squares [divisor = $u(\$12)$] are above the solid diamonds [divisor = $u(\$24)$], which are above the symbols for the larger divisors. This pattern of two rank orders has been observed when subjects are asked to judge "ratios" and "differences" of stimulus intervals or distances, but it is not typically obtained when subjects are asked to judge "ratios" and "differences" of stimulus magnitudes (Birnbau, 1978, 1982; Birnbau, Anderson, & Hyman, 1989; Hagerty & Birnbau, 1978).

For many psychophysical continua, judgments of "ratios" and "differences" do not obey the predictions in Fig. 3, but instead can be represented by the theory that the subject evaluates differences in both cases, but there are different judgment functions in the two tasks (Birnbau, 1978, 1980, 1982; Birnbau & Elmasian, 1977; Birnbau & Mellers, 1978; Birnbau & Veit, 1974; Elmasian & Birnbau, 1984; Hardin & Birnbau, 1990; Mellers, Davis, & Birnbau, 1984; Veit, 1978). Subtractive theory replaces Eq. (3) with the difference model for "ratios":

$$R_{xy} = J_R[u(x) - u(y)]. \quad (5)$$

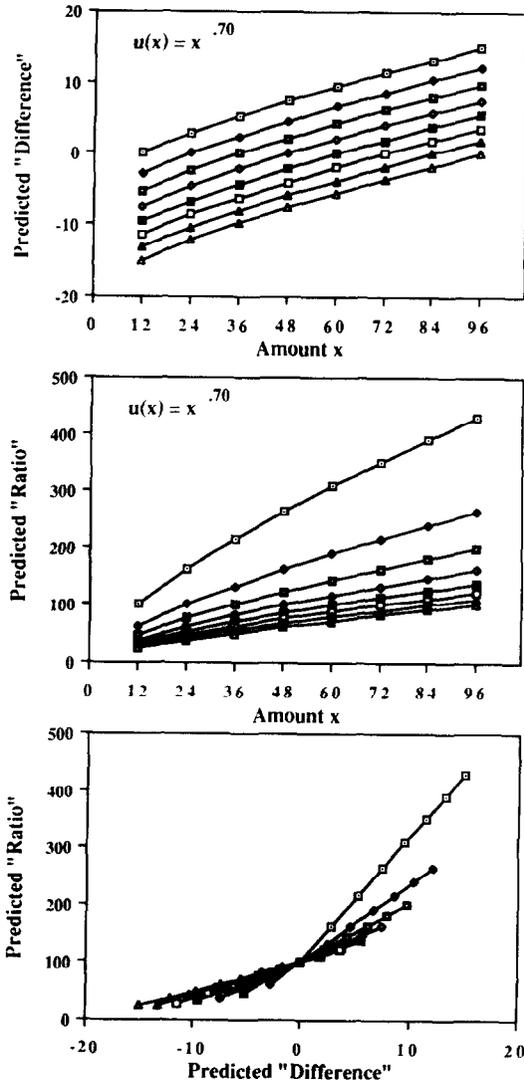


FIG. 3. Predicted "ratios" and "differences" according to the theory that subjects use the two operations (as instructed), illustrated with $u(x) = x^{.70}$. (Top) Predicted "differences" plotted against Amount x with a separate curve for each level of Amount y . (Center) Predicted "ratios," plotted in the same fashion. (Bottom) Predicted "ratios" plotted against "differences."

This theory implies that "ratios" and "differences" are monotonically related. Should this result occur, then the subtractive theory can be fit to the data, yielding a scale of utility that is unique (in principle) to an interval scale.

Figure 4 shows predictions for this subtractive theory, plotted in the same fashion as Fig. 3, again using $u(x) = x^{.70}$. "Ratios" are an exponential function of subjective differences [the base was set to 1.1175 to put "ratios" on a scale from 1/8 to 8, and they have been multiplied by 100]. Note that predicted "ratios" diverge, predicted "differences" are paral-

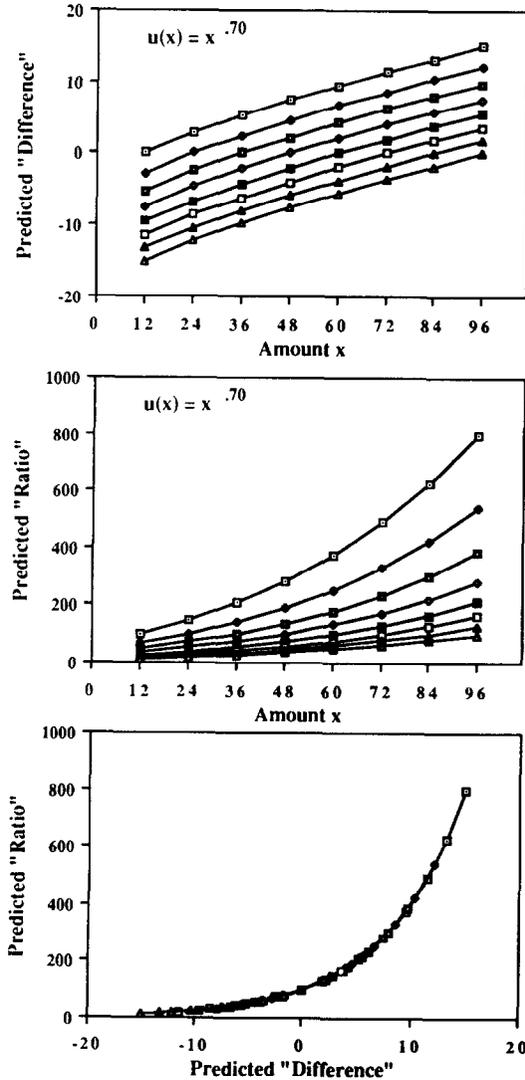


FIG. 4. Predicted "ratios" and "differences" according to the theory that both tasks are governed by subtraction, illustrated using $u(x) = x^{.70}$. Panels plotted as in the legend to Fig. 3.

lel, and predicted "ratios" are a monotonic function of predicted "differences."

The fact that so many psychophysical continua seem to obey one-operation theory has led to the hypothesis that many continua are inherently interval scales, on which "ratios" and "differences" of intervals are meaningful, but simple "ratios" and "differences" are not. It seems plausible that monetary increments might be perceived in either way, as designating wealth positions or increments from the status quo, and, indeed, these two views led to two different interpretations of the utility function itself (Edwards, 1954).

One operation and two operation theory can be compared by asking whether "ratios" and "differences" are related as in the lower panel of either Fig. 3 or Fig. 4. Furthermore, if subjects use a ratio operation for "ratios" and if $u(x)$ is any power function of x , then judged "ratios" should be a strictly monotonic function of actual ratios of monetary amounts. In sum, if either the two operation theory or the one operation theory is correct, scales can be derived that are unique to at least an interval scale. In either case, these scales can be compared with the scales derived from models of the buyer's and seller's prices.

METHOD

Subjects performed four tasks: in two tasks, they judged the value of gambles from the buyer's and seller's points of view, and they also judged "ratios" and "differences" of subjective magnitudes of the values of monetary amounts. Subjects were given a brief overview of all tasks at the beginning, and the order of performing the tasks was counterbalanced across subjects.

Instructions: "Ratio" and "Difference" Tasks

Subjects were instructed that they would be judging the psychological values of different amounts of money in two separate ways: by judging the "ratio" of the subjective magnitudes and the "difference" in the subjective magnitudes of the two amounts of money.

Each trial was presented using the format of the following example:

\$12 \$48

The "ratio" task called for judgments of the preference for the first amount of money relative to the second amount, using a modulus of 100. Seven examples were as follows:

- 800 = prefer the first amount 8 times more than the second,
- 400 = prefer the first amount 4 times more than the second,
- 200 = prefer the first amount 2 times more than the second,
- 100 = prefer the two amounts equally,

- 50 = prefer the first amount $\frac{1}{2}$ as much as the second,
 25 = prefer the first amount $\frac{1}{4}$ as much as the second,
 12.5 = prefer the first amount $\frac{1}{8}$ as much as the second.

Subjects were encouraged to use any numerical values, more extreme or in between, consistent with this scheme, that best represented their "psychological ratios of preference."

In the "difference" task, subjects rated the subjective "difference" between receiving the first and the second amount of money, using a scale with category labels varying from -15 ("prefer the second amount *very very much more* than the first") to 15 ("prefer the first amount *very very much more* than the second"), with 0 being "equal preference for the two amounts."

Instructions: Buyer's and Seller's Tasks

The gambles were described as lotteries in which a slip of paper would be drawn from a bin containing 100 slips with monetary amounts printed on them (with frequencies varying in proportion to specified probabilities), and the amount printed on the slip would be paid to the player. Each lottery was displayed using the format of the following example:

.20	.80
\$12	\$48

which represented a gamble with a .20 chance to win \$12, and a .80 chance to win \$48, described as 20 slips per 100 with \$12 and 80 slips with \$48.

Point of view instructions read (in part) as follows:

"In the *buyer's point of view*, imagine that you are deciding the *most that a buyer should pay* to buy the chance to play the lottery. The buyer exchanges money for the chance to play the lottery. If the lottery pays off, the buyer can profit; if not, the buyer can lose the difference between the amount won and the price of the lottery. Therefore, the buyer wants to *pay as little* as possible and to have as good a chance of winning good payoffs as possible . . . The buyer will be upset with you if you advise him to pay for gambles that lose too much money. Remember, you are asked to advise the buyer the *most* he should be willing to pay to get to play the lottery."

"In the *seller's point of view*, imagine that the seller owns the lottery ticket and you are deciding what is the *least that a seller should accept to sell the lottery*. The seller receives money and gives up the chance to play the lottery. Thus, if the lottery pays off, the seller would have been better off to keep the lottery than to have sold it, but if the lottery doesn't pay off (or pays only a little), the seller profits by selling. Therefore, the seller

wants to *receive as much* as possible for each gamble . . . The seller will be upset if you advise him to sell a gamble for a low price that pays off the buyer a lot of money. Remember, you are asked to advise the seller concerning the *least* the seller should accept to *sell* the lottery, rather than play it himself.”

Designs

For the “ratio” and “difference” tasks, the levels were the following: \$0, \$12, \$24, \$36, \$48, \$60, \$72, \$84, and \$96. Stimulus pairs were generated from a 9 by 8 (Amount x by Amount y) factorial design, in which only the first factor included the level \$0, and the second factor included the other eight levels.

In the gamble tasks, the stimuli were generated from three designs. The first was an Amount x by Amount y , triangular design in which the first amount was always less than the second, and the same nine values from \$0 to \$96 were used, generating 36 pairs. Each gamble in this design consisted of a 50–50 chance of winning either Amount x or Amount y . This design was intended to permit a test of configural weight theory and allow estimation of $u(x)$ for comparison with the “ratio” and “difference” tasks.

The second design was a 4×7 , Amount Pair by Probability, factorial. The first factor determined the two monetary amounts that could be won (\$0 vs \$72, \$0 vs \$96, \$24 vs \$72, or \$24 vs \$96), and the second factor was the probability of winning the larger of the two amounts rather than the smaller (.05, .1, .2, .5, .8, .9, .95). This design was used to replicate the violations of monotonicity observed by Birnbaum *et al.* (1992).

The third design consisted of 16, 50–50 gambles between amounts that differed by \$6, from \$0 to \$96: (\$0, \$6), (\$6, \$12), . . . (\$90, \$96). It was expected that these lotteries would be minimally affected by point of view; in combination with other lotteries that would be affected by point of view, these trials were used to demonstrate that the rank order of lotteries changed as a function of point of view.

Procedure

The stimulus pairs in the “ratio” and “difference” tasks and the gambles in the “buyer’s-point-of-view” and the “seller’s-point-of-view” tasks were printed in random order in individual booklets. The 80 trials from the three gamble designs were intermixed and printed in random order as well. There were 4 separate booklets containing the instructions, 10 representative warm-up trials for each task, and the randomized experimental trials. These were given to the subjects in one of eight orders, with either 12 or 13 subjects in each of the following: DRBS, RSDB, RDSB, DBRS, BSDR, SRDB, SBRD, and BDSR, where B = “buyer’s-

point-of-view," S = "seller's-point-of-view," R = "ratio" and D = "difference" tasks, respectively. Each subject finished each task at his or her own pace before proceeding to the next. Most subjects completed the tasks in approximately 1½ hr.

Subjects

The subjects were 100 California State University, Fullerton undergraduates who participated for extra credit in Introductory Psychology.

Choice Task

To check whether subjects would violate monotonicity when offered a choice between gambles, 72 of the subjects were also asked to choose which gamble they would prefer to play for each of 8 pairs of gambles. Included fifth was the choice between (\$96, .95, \$0) and (\$96, .95, \$24). The gambles were presented in the same format as in the buyer's and seller's tasks, except the two gambles in each pair were displayed side by side, and the subject was instructed to circle the preferred gamble. This task was always presented last. [The other seven choices were as follows: trials 3, 1, and 7 were choices between (\$96, .5, \$0) and (\$18, .5, \$24), (\$30, .5, \$36), and (\$42, .5, \$48), respectively; trials 2 and 4 offered choices between three-outcome gambles with .1 chance to win \$5, .4 to win \$16, and .5 to win y , vs (\$8, .5, y), in which y was either \$0 or \$80, respectively; trials 6 and 8 offered choices between (\$12, .5, x) and three-outcome gambles of .2 to win \$1, .3 to win \$24 and .5 to win x , in which x was either \$0 or \$96.]

RESULTS

Buying and Selling Prices

Figure 5 shows the median buying and selling prices of ($x.5y$) gambles, i.e., 50–50 chances to win either x or y , plotted as a function of Amount x , with a separate curve for each level of Amount y .³ Buying and selling prices are quite different in three respects: (a) Buying prices are lower than selling prices, with the greatest differences occurring for largest ranges ($|x - y|$) of value. (b) Buying prices diverge to the right, whereas selling prices are nearly parallel; and (c) the rank order of the two judgments is quite different.

The data in Fig. 5 resemble the predictions of configural weight theory shown in Fig. 1, which show changes in the shape of the curves and rank

³ A 50–50 gamble between x and y is a sure thing to win x when $x = y$; therefore, the value x is plotted in Fig. 5 for these cases (even though they weren't presented for judgment), to facilitate comparison with Fig. 1.

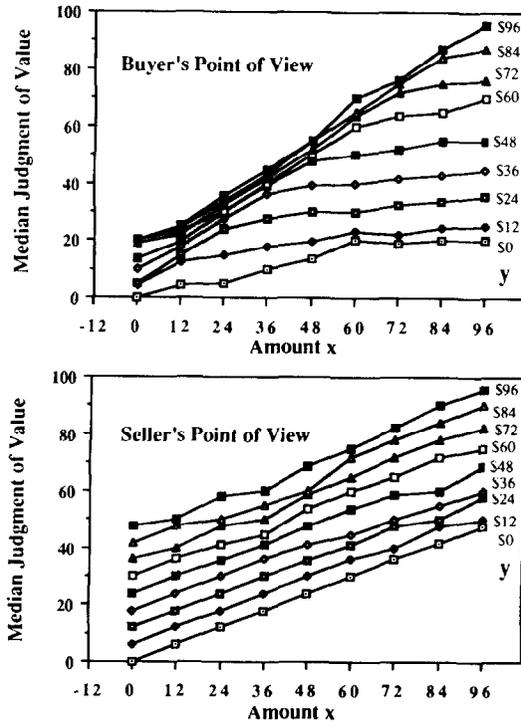


FIG. 5. Median buyer's prices (top panel) and seller's prices (bottom panel), plotted as in the legend to Fig. 1.

order of the points that are due to changes in the configural weight parameter with the u function fixed. Mean judgments and judgments of individual subjects, plotted in the same fashion as the medians in Fig. 5, showed similar patterns.⁴

⁴ For the present data, different measures of central tendency would yield the same conclusions. Means and medians of judgment data are usually monotonically related, but medians have two properties that make them theoretically attractive for buyer's and seller's prices. First, medians are less affected by outliers than means, and because half of the subjects would vote to pay more and half to pay less, medians seem an attractive compromise for a monetary amount (Galton, 1907). Second, the rank order of medians (the smallest value for which the cumulative proportion of responses is greater than or equal to .5) is invariant under monotonic transformation of the data. For "ratios" and "differences," the theory that is most compatible with previous data indicates that geometric means and arithmetic means are good aggregate measures, respectively. "Ratio" judgments are theorized to be exponentially related to subjective values (Birnbbaum, 1980, 1982) and have been found to be linearly related after logarithmic transformation (see e.g., Hardin & Birnbbaum, 1990). Geometric means (the antilogs of the mean logs of the responses) are also typically quite close to medians for "ratio" judgments.

Some subjects displayed even more extreme changes, as in Fig. 6, which presents the individual judgments of the first subject, whose buying and selling prices changed from divergence to convergence. Such a subject gives the appearance of changing from “risk aversion” to “risk seeking” as the point of view is changed, since the buying prices are less than expected value and the selling prices exceed expected value; nevertheless, such a pattern is compatible with a single utility function according to configural weighting theory (Fig. 1). It is interesting to note that this subject reported that she purchased state lottery tickets each week. Such behavior is also compatible with “risk aversion” in configural weighting theory (Fig. 2).

The data were analyzed by two monotonic transformation procedures, in order to test the implications of the SEU type models [Eq. (1)] against the configural-weight models. The first analysis used a modified version of MONANOVA (Kruskal & Carmone, 1969) to find M^{-1} to optimize the fit of Eq. (1) to the median buyer’s and seller’s prices. Figure 7 shows that the utility scales estimated from buyer’s price judgments are concave downwards relative to the utility scales estimated from seller’s prices.

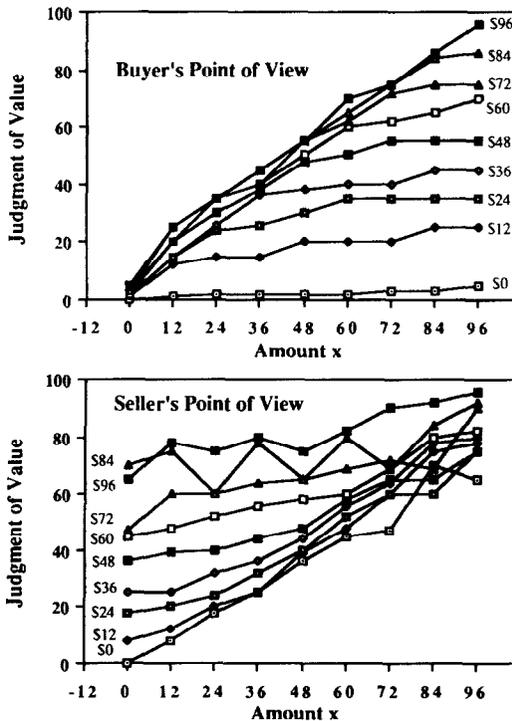


FIG. 6. Buyer's and seller's prices for one subject, as in the legend to Fig. 5.

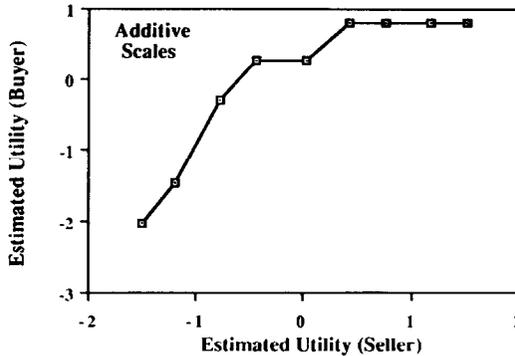


FIG. 7. Comparisons of scale values estimated from additive model. Scales estimated from additive model are nonlinearly related, consistent with prediction of configural weighting theory.

Such downward curvature is predicted by configural-weight theory if the $u(x)$ function is invariant and W_V is greater for seller's than for buyer's prices (Birnbäum *et al.*, 1992), but the additive models would require different $u(x)$ functions for buyer's and seller's prices, as shown in Fig. 7.

The second set of analyses (of the same data) used MONANOVA to fit Eq. (1) separately for the two cases in which $x < y$ and where $x > y$. This procedure yields four estimates of the utility function, for buyer's and seller's prices, and for the lower and higher amounts. These four estimated utility functions are plotted against the scales derived from all of the data in Fig. 8.

Figure 8 shows that when the buyer's prices are analyzed as an additive function of the higher and lower values (i.e., for $x < y$ only), the estimated scales for the lower and higher amounts are nonlinearly related and have different slopes compared to the scale estimated from the buyer's prices fit to the additive model over the entire factorial domain. This pattern is predicted by configural weight theory, if $W_V < 0$, but it is inconsistent with the additive model, since Eq. (1) would imply that the estimated utility function should be the same, whether it is estimated separately for the lower and higher amounts or estimated simultaneously for all amounts.

For seller's price medians, however, the scales estimated separately for higher and lower values were nearly linearly related to the additive scales, and nearly parallel, as shown in the right of Fig. 8, consistent with the idea that W_V is near zero for the median seller's judgments.

Dominance Violations

Figure 9 shows median buying and selling prices plotted as a function of the probability of winning the larger amount (either \$72 or \$96), with a

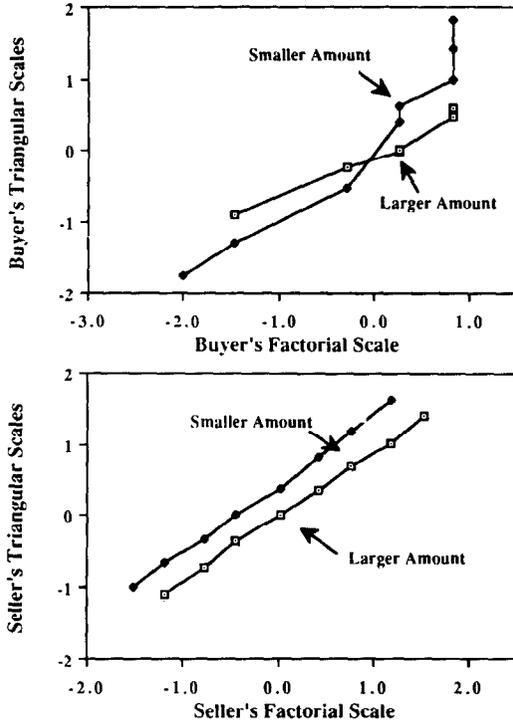


FIG. 8. Scales of utility for x and y estimated from the additive model for $x < y$ plotted against scale values estimated from entire domain. (Top) Buyer's prices. (Bottom) Seller's prices.

separate curve for each level of the smaller amount (\$0 or \$24). These data replicate the trends found by Birnbaum *et al.* (1992) and the pattern of predictions in Fig. 2: median prices can be higher for dominated gambles when the outcome of receiving \$0 has a small probability.

For each subject, eight tests of dominance for $x = \$0$ vs \$24 with $p = .9$ and $p = .95$ were counted for $y = \$72$ and \$96, for buyer's and seller's judgments. Ties were not counted as violations. (Seven of these eight comparisons of medians violated dominance). Out of eight comparisons, 63 of 100 subjects had three or more violations and 45 subjects had four or more violations. Only eight subjects had no violations of dominance. In the choice task, however, when (\$96, .95, \$0) was directly compared with (\$96, .95, \$24), only 5 of the 72 subjects violated dominance.

Consistent with the theory that the stimulus \$0 has lower weight, offered to explain Fig. 9, the buyer's prices in Fig. 5 show a steeper slope for the lowest curve ($y = \$0$), for the cases in which $x > y$ (this statement applies only to the right most portion of each curve, where $x > y$). If \$0

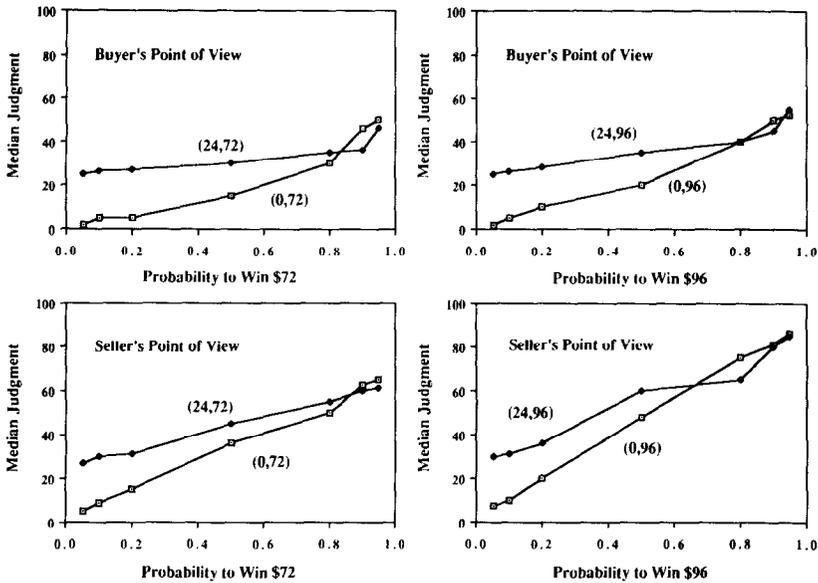


FIG. 9. Median buyer's and seller's prices plotted as a function of probability of receiving either \$96 (right panels) or \$72 (left panels), with a separate curve for each value of the lower outcome (\$0, or \$24). Solid diamonds represent gambles that dominate open squares. Crossing of curves indicate violations of dominance.

has a lower weight and the weights sum to one, then the relative weight of x as the larger value would be increased when paired with $y = \$0$.

"Ratio" and "Difference" Judgments

Geometric mean "ratio" judgments and mean "difference" judgments are shown in Fig. 10. Data are plotted as a function of Amount x , with a separate curve for each level of Amount y , as in Figs. 3 and 4. Geometric mean "ratios" are plotted against mean "differences" in the lower panel of Fig. 10. The two judgments are neither monotonically related, as in Fig. 4, nor do they quite match the pattern of two operations shown in Fig. 3.

Individual data were examined, and were sorted into three categories according to their relationships between "ratios" and "differences."⁵ (a) There were 25 subjects who used a small number of distinct responses for either one task or the other. For example, one such pattern would occur

⁵ Subjects were assigned to this category on the basis of the union of two criteria. If (a) the majority of $[R(84,12), R(72,12), R(60,12)]$ were greater than $R(96,24)$ and the majority of $[D(84,12), D(72,12), D(60,12)]$ were less than or equal to $D(96,24)$ or (b) if the majority of $[R(72,12), R(60,12), R(48,12)]$ were greater than $R(96,36)$ and the majority of $[D(72,12), D(60,12), D(48,12)]$ were less than or equal to $D(96,36)$, then the subjects were placed in this category. Note that these criteria distinguish the predictions in Figs. 3 vs 4.

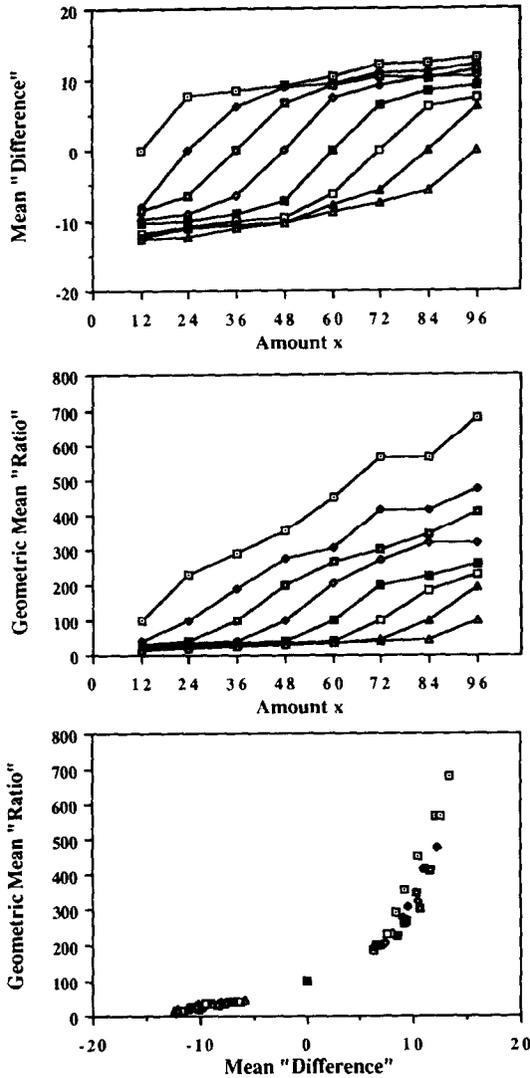


FIG. 10. Geometric mean "ratios" and mean "differences," plotted as in the legends to Figs. 3 and 4.

if a subject said it was always "8 times" more preferable to receive the larger amount of money in any pair of values and " $\frac{1}{8}$ " as preferable to receive the smaller amount. Although such data are compatible with one-operation theory, these data are not really diagnostic. (b) Some subjects appeared to have evidence of two operations, and a couple even calculated the ratios exactly. Individual judgments for one subject of the 29

subjects with evidence of two operations are shown in Fig. 11. (c) The remaining 46 subjects were averaged, and their data appeared well-approximated by the theory that they used only one operation for both tasks, as in Fig. 4.

Figure 12 plots mean log "ratio" judgments as a function of the log of

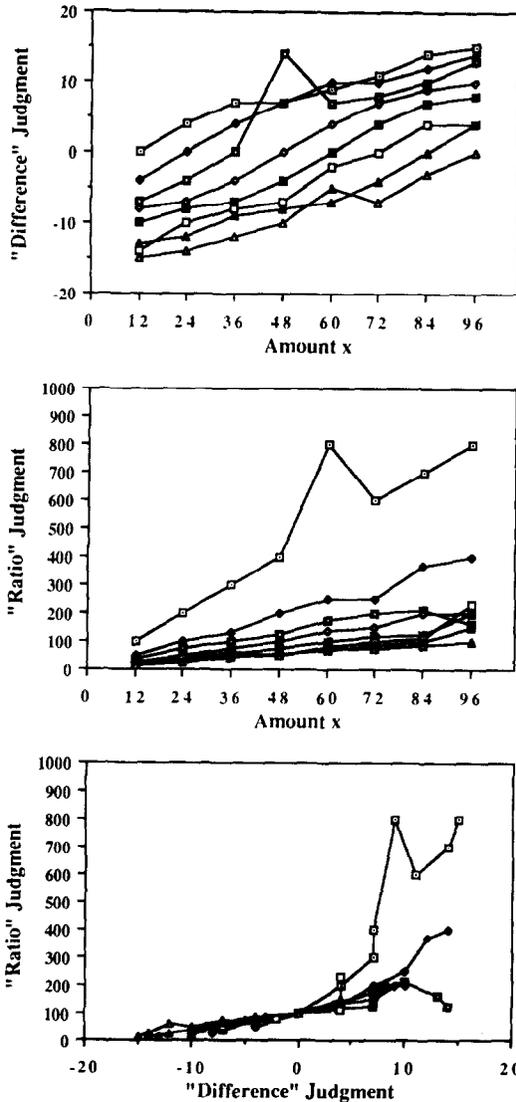


FIG. 11. Individual judgments of "ratios" and "differences" for one subject who appeared to use two operations, as in the legend to Fig. 3.

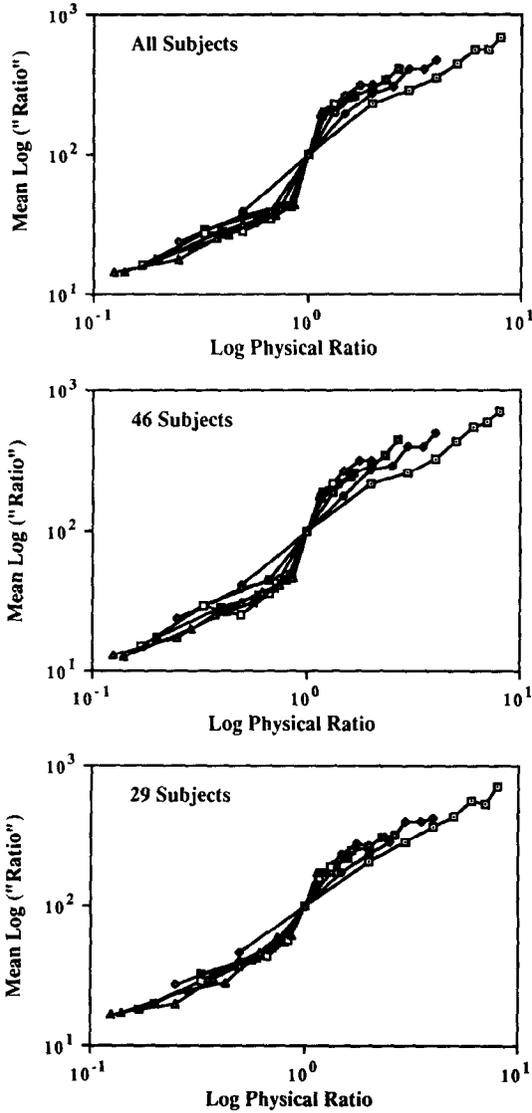


FIG. 12. Geometric mean "ratios" as a function of physical ratios of monetary amounts for all 100 subjects (top), 46 subjects who used the full range of responses but did not show evidence of two operations (center), and 29 subjects who showed evidence of two operations (bottom), plotted in log-log coordinates. Curves differ systematically from a single function, indicating evidence against the assumption that "ratio" judgments are a ratio of power functions.

the physical ratio of monetary amounts, (x/y) , with separate curves for values of y , and a separate panel for each group of subjects. According to the ratio model, any power function for u implies that "ratios" should be a monotonic function of physical ratios, because if $u(x) = x^b$, and if $R_{xy} = J_R[u(x)/u(y)]$, then $R_{xy} = J_R[x^b/y^b]$; hence $R_{xy} = J_R[(x/y)^b]$; therefore, $R_{xy} = J_R^*[x/y]$, where J_R^* is strictly monotonic. Figure 12 shows that judged "ratios" systematically deviate from the pattern predicted from a ratio of power functions. Although the data for the 29 subjects who showed signs of two-operations come closer to this prediction, even those data show the same pattern of deviation: as a given physical ratio is moved up the scale (e.g., 24/12, 48/24, 96/48), the average "ratio" judgment increases. If "ratio" judgments were governed by subtraction, however, then the results in Fig. 12 would not contradict the power function.

Model Analyses

The data of Figs. 5 and 10 were fit to the following model:

$$R_{xy} = J_R[u(x) - u(y)], \quad (6)$$

$$D_{xy} = J_D[u(x) - u(y)], \quad (7)$$

$$B_{xy} = c[.5(u(x) + u(y)) + W_B|u(x) - u(y)|]^a, \quad (8)$$

$$S_{xy} = c[.5(u(x) + u(y)) + W_S|u(x) - u(y)|]^a, \quad (9)$$

where R_{xy} and D_{xy} are geometric mean "ratios" and mean "differences" between the utilities of x and y ; B_{xy} and S_{xy} are median buying and selling prices for 50-50 gambles to win either x or y , respectively. The constants, c and a in Eqs. (8 and 9), allow for judgmental effects and for u^{-1} in Eq. (2).

The "ratios" and "differences" were fit to Eqs. (6 and 7) separately, using MONANOVA to estimate monotonic transformations, J_R^{-1} and J_D^{-1} , so as to transform the data to fit the subtractive model. These transformations define new dependent variables, as follows: $r_{xy} = J_R^{-1}(R_{xy})$ and $d_{xy} = J_D^{-1}(D_{xy})$, where r_{xy} and d_{xy} are the transformed values that preserve the rank order of the data.

The transformed values were then fit, along with the buyer's and seller's medians, by means of a special computer program, SARFIT, that was written to find a single scale, $u(x)$, so as to minimize a weighted sum of squares across the four data arrays. Within each array, the sum of squared deviations between data and predictions was calculated, and this sum was divided by the sum of squared deviations for the data about the mean within that matrix (see Birnbaum, 1980).

A separate value of $u(x)$ was permitted for each of the eight levels of x , but $u(\$12)$ was arbitrarily set to 1.0, leaving seven parameters to estimate for the levels of money. The function converting utility of a gamble to a

monetary judgment was estimated as a power function and multiplied by a constant (c) to put the response in dollars; two configural weights were estimated for the two points of view; therefore, there were 11 parameters to estimate, not counting the monotonic transformations made by MONANOVA.

The estimated utility values were as follows: 1.0, 3.21, 4.60, 6.41, 8.35, 9.99, 11.13, and 12.24 for the levels from \$12 to \$96, respectively. The estimated values of W_v were $-.266$ and $.011$ for buyer's and seller's points of view, respectively. The exponent and multiplier in Eqs. (8 and 9) were $a = 1.026$ and $c = 6.885$, respectively.

The predictions of Eqs. (6–9) based on these parameters give a fairly good approximation to the data. For example, the predicted buyer's price for (\$12, .5, \$96) would be calculated from Eq. (8) as follows:

$$6.885(.5(1.0 + 12.24) - .266(11.24))^{1.026} = 25.9,$$

which is close to the observed median of 25. The percentages of deviations for r_{xy} , d_{xy} , B_{xy} , and S_{xy} were .49, .41, 1.40, and 1.58% of the variance within each array, respectively. Deviations from the model appear attributable to two main sources: (a) the predictions for "ratios" and "differences" are monotonically related, whereas the average data show some slight but systematic deviations (see bottom panel of Fig. 10). (b) The power function is not optimal to represent the transformation from subjective value to response in Eqs. (8 and 9).

DISCUSSION

Buying and Selling Prices

The present buying and selling prices are compatible with previous results obtained by Birnbaum *et al.* (1992); they agree with Birnbaum and Stegner's (1979) theory of point of view; and the violations of dominance appear consistent with Birnbaum *et al.* (1992) theory of the weighting of \$0 outcomes.

These data show that the preference order can be altered in two ways: (a) By changing from buyer's to seller's points of view, the rank order of judgments can be changed in the manner predicted by configural weight theory. These changes can be explained by the assumptions that the $u(x)$ function is invariant and the W_v parameter changes, consistent with the theory of Birnbaum *et al.* (1992). (b) In direct choice, few subjects choose the dominated gamble; however, judgments of the buying or selling prices of gambles involving an outcome of \$0, stated on separate occasions, can show predictable violations of dominance. These violations can be explained by postulating that the weight of \$0 behaves differently as a func-

tion of probability than does the weight of positive outcomes (for the judgment tasks but not necessarily for the choice tasks).

Preference Reversals

These changes in preference order may seem reminiscent of the relation between bids and choices discussed by Lichtenstein and Slovic (1971), Slovic and Lichtenstein (1983), Goldstein and Einhorn (1987), Grether and Plott (1979), and Tversky, Sattath, and Slovic (1988), among others. However, the present results represent distinct phenomena. Those authors found that among gambles with roughly equal expected value, subjects would set higher value on gambles with higher amounts to win, but when given a choice, people prefer gambles with higher probabilities to win. Tversky *et al.* (1988) theorized that probability has a greater weight relative to amounts to win in choice tasks, whereas amounts to win have greater relative weight in pricing tasks.

Preference reversals among probability by outcome combinations are distinct from the present phenomena, which cannot be explained or predicted by the theories proposed in the studies cited above. Those explanations refer to changes in the way that the probability of an outcome and the value of that outcome combine. In contrast, the present phenomena involve the way in which different outcomes combine: changes in rank order can be observed even when the level of probability is held constant, as in Fig. 5, which displays only 50–50 gambles.

The violations of monotonicity (dominance) which occur for pricing judgments (but not choice) represent another distinct type of preference reversal. Although they are observed by comparing pricing with choice, they also involve the manner in which outcomes combine, rather than probability and outcome. These violations of monotonicity are explained by the assumption that the outcome of zero receives lower weighting at each level of probability than the weighting of equally probable nonzero outcomes.

Conversely, it appears from other findings that neither configural weighting nor simple contingent weighting alone can explain the preference reversals between ratings and bids. Mellers, Ordóñez, and Birnbaum (1992) found that differences in rank order between ratings and bids are not consistent with the theory that only the weights changed, but instead can be explained by the theory that probability and amounts combine by multiplication for bids, while the same variables combine by an additive process for ratings. Furthermore, they concluded that the process by which these variables are combined in ratings could be altered by changing the set of stimuli to include near zero, zero, and negative values.

“Ratios” and “Differences”

Although the results of the buyer's and seller's price judgments are

compatible with previous research, the results of the experiment on "ratios" and "differences" seem more puzzling. Previous research has led to two distinct cases, where "ratios" and "differences" appear to be governed by one operation or by two. When asked to judge "ratios" and "differences" of loudness, heaviness, pitch, likeableness, darkness, prestige of occupations, easterliness, or westerliness of US cities, the data have been remarkably consistent with the conclusion that subjects use the subtractive operation to compare stimuli in both tasks (Birnbbaum, 1978, 1980, 1982; Hardin & Birnbbaum, 1990; Veit, 1978). However, when subjects are asked to judge "ratios of differences" and "differences of differences" or "ratios of distances" and "differences of distances," data have conformed to the theory that both ratios and differences operate on a common scale of distances (Birnbbaum, 1978, 1982; Birnbbaum *et al.*, 1989; Veit, 1978).

The present data show greater individual differences on the "ratio"/"difference" question in comparison with previous research. Furthermore, even when the subjects who showed evidence of two operations are segregated, the data do not appear entirely consistent with two operation theory. Perhaps because the stimuli are presented in numerical form, the tasks invite numerical operations, so some subjects on some occasions may have actually computed numerical ratios and differences. Indeed, the stimuli were integer multiples of 12, with many "easy" 2:1 ratios (24/12; 48/24; 96/48), and the largest example "ratio" response ("8") matched the largest physical ratio ($96/12 = 8$).

Nevertheless, "ratio" judgments were not a monotonic function of physical ratios, even for the subjects whose data most resembled the predictions of two operation theory. The failure of "ratio" judgments to be a function of physical ratios constitutes evidence against the power function, assuming the ratio model.

Is Utility a Power Function of Money?

A number of authors have theorized that the utility function for money might be represented as a power function (Galanter, 1962; Stevens, 1975; Breault, 1983; Parker, Schneider, Stein, Popper, Darte, & Needel, 1981). Furthermore, when monetary amounts are presented as numbers, it seems plausible that money scales might be consistent with scales of the psychological magnitudes of numbers.

Rule and Curtis (1982) summarized a number of studies of the subjective magnitudes of numbers (using a variety of methods) that were fit with power functions with exponents between .4 and 1, with an average between .6 and .80. Birnbbaum's (1974b) scale, derived from contextual effects in ratings, and Rose and Birnbbaum's (1975) scale, derived from

nonnumerical "ratio" and "difference" judgments, were also well-approximated by values in this range.

There is a more compelling argument for a power function for money, based on regularities for gambles between \$0 and x . A number of studies have investigated such gambles, and data appear consistent with the assumption that when the response is a price, the data have two properties that provide nontrivial support for the power function. The argument goes as follows: Suppose the subjective value of gambles with a probability to receive x and otherwise receive nothing can be represented as the product of some function of probability and another function of x :

$$g_{px} = s(p)u(x), \quad (10)$$

where g_{px} is the subjective value of the gamble; $s(p)$ is a function of probability (and may include a configural weight); and $u(x)$ is the utility function for money, as before. If $u(\$0) = 0$, Eq. (10) follows from virtually all utility theories, including the SEU-type theories, configural-weight theory, and Tversky *et al.*'s (1988) contingent-weight theory. When the subject gives a monetary response, G_{px} , we can represent the judgment as follows:

$$G_{px} = J[g_{px}], \quad (11)$$

where J is the judgment function that maps subjective value into a monetary response. If J is a linear function of any power function, then the data should form a bilinear fan in which the response for any level of p should be a linear function of the response for another level of p . This property has been tested by many experimenters and has generally been successful (Anderson & Shanteau, 1970; Birnbaum & Mellers, 1988; Komorita, 1964; Shanteau, 1974, 1975; Tversky, 1967a,b). Birnbaum (1983) presents another example of a reasonable multiplicative model in which the dependent variable is a monetary scale.

Furthermore, when the response is in money, as well as the stimulus, it is inviting to theorize that $J = u^{-1}$, and if $u(x) = x^b$, we can compose Eqs. (10 and 11) as follows:

$$G_{px} = (s(p)x^b)^{1/b}, \quad (12)$$

therefore, defining $S(p) = [s(p)]^{1/b}$,

$$G_{px} = S(p)x. \quad (13)$$

Therefore, if $u(x)$ is a power function of x and if the response function is the inverse of u , it follows that bids for such gambles should be directly proportional to x . This property has been tested by Birnbaum and Mellers (1989), and has been found to be a close approximation to the data (see

Mellers, *et al.*, 1992). Such findings make it inviting to retain the power function until we are forced to give it up.

We then return to the finding that "ratios" of utility are not a function of physical ratios. If $u(x) = x^b$, and if subjects use a ratio operation on $u(x)$, then "ratios" should be a monotonic function of physical ratios, contrary to Fig. 12. Similar violations of the power function have occurred for loudness of tones (Birnbau & Elmasian, 1977) and heaviness of lifted weights (Mellers *et al.*, 1984), and those studies also concluded that the ratio model and power function are not compatible. However, in those studies, "ratios" and "differences" were monotonically related, and the authors represented "ratio" judgments with the subtractive model. In the present study, the data show trends that indicate at least some subjects some of the time used two operations. In psychophysics, the power function has historical appeal, but the arguments for it, aside from its flexibility in approximating empirical curves, are not compelling. It is perhaps ironic that "ratio" judgments are incompatible with the ratio model of power functions, because it was the use of "ratio" judgments and the assumption of a ratio operation that led Stevens (1975) to the power function.

Conclusions

We are left with several possibilities to account for our data: Individual differences, especially for "ratio" judgments, may obscure the proper estimation of the $u(x)$ functions and proper tests of scale convergence among the tasks. It would be desirable to conduct the studies to gather sufficient data at the individual level to separately estimate a stable $u(x)$ function for each subject in each array. It may be that the request to judge "ratios" and "differences" of monetary amounts presented numerically may have induced in some subjects on some trials an intellectual approach in which some calculated ratios were mixed with subjective differences. Such a mixture might lead to the "in-between" pattern of data that neither conform to the predictions in Fig. 3 nor those in Fig. 4. Perhaps such tendencies could be avoided by asking subjects to compare utilities for monetary amount with utilities for goods. Such an approach might reduce the likelihood that subjects would mix strategies in the experiment.

If we are willing to disregard the minority of subjects who appear to be using two operations (perhaps to suspect them of using numerical calculation only part of the time), the data form a simpler picture that does allow a rough approximation of all of the data with a single scale. We can approximate most of the data by assuming one operation theory [Eqs. (6 and 7)] for "ratios" and "differences" and configural weight theory [Eqs. (8 and 9)], with different values of W for the different points of view.

Although this representation does not work perfectly, it does provide a reasonable summary of four arrays of data using a single scale of utility.

Allowing a different scale of utility in each array would provide a better fit for Eqs. (6–9); however, the inconsistency within the buyer's prices (Fig. 8) indicates that transformation theories that try to explain the results without configural weighting do not give a consistent account of the data. With configural weighting, the data encourage the belief that the phenomena that led to different utility and value functions can be explained with a single utility function.

REFERENCES

- Allais, M. (1979). The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the American School. In M. Allais & O. Hagen (Eds.), *Expected utility hypothesis and the Allais paradox*. Dordrecht: Reidel.
- Anderson, N. H., & Shanteau, J. (1970). Information integration in risky decision making. *Journal of Experimental Psychology*, *84*, 441–451.
- Becker, J. & Sarin, R. K. (1987). Lottery dependent utility. *Management Science*, *33*, 1367–1382.
- Bell, D. E., & Raiffa, H. (1988). Marginal value and intrinsic risk aversion. In D. Bell, H. Raiffa, & A. Tversky (Eds.), *Decision-making: Descriptive, normative, & prescriptive interactions*. New York: Wiley.
- Birnbaum, M. H. (1974a). The nonadditivity of personality impressions. *Journal of Experimental Psychology Monograph*, *102*, 543–561.
- Birnbaum, M. H. (1974b). Using contextual effects to derive psychophysical scales. *Perception & Psychophysics*, *15*, 89–96.
- Birnbaum, M. H. (1978). Differences and ratios in psychological measurement. In N. J. Castellan & F. Restle, (Eds.), *Cognitive Theory*, (Vol. 3 pp. 33–74). Hillsdale, NJ: Erlbaum.
- Birnbaum, M. H. (1980). Comparison of two theories of "ratio" and "difference" judgments. *Journal of Experimental Psychology: General*, *109*, 304–319.
- Birnbaum, M. H. (1981). Reason to avoid triangular designs in nonmetric scaling. *Perception & Psychophysics*, *29*, 291–293.
- Birnbaum, M. H. (1982). Controversies in psychological measurement. In B. Wegener (Ed.), *Social attitudes & psychophysical measurement*. (pp. 401–485). Hillsdale, NJ: Erlbaum.
- Birnbaum, M. H. (1983). Perceived equity of salary policies. *Journal of Applied Psychology*, *68*, 49–59.
- Birnbaum, M. H. (1987). Searching for coherence in judgment and decision making. Invited address to Western Psychological Association Meetings, Long Beach. (Available from Conference Coordinating Company, 37 Indian Rock Road, San Anseimo, CA 94960)
- Birnbaum, M. H., Anderson, C. A., & Hynan, L. G. (1989). Two operations for "ratios" and "differences" of distances on the mental map. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 785–796.
- Birnbaum, M. H., Coffey, G., Mellers, B.A., & Weiss, R. (1992). Utility Measurement: Configural weight theory and the judge's point of view. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 331–346.
- Birnbaum, M. H., & Elmasian, R. (1977). Loudness "ratios" and "differences" involve the same psychophysical operation. *Perception & Psychophysics*, *22*, 383–391.
- Birnbaum, M. H., & Jou, J.-W. (1990). A theory of comparative response times and "difference" judgments. *Cognitive Psychology*, *22*, 184–210.

- Birnbaum, M. H., & Mellers, B. A. (1978). Measurement and the mental map. *Perception & Psychophysics*, *23*, 403-408.
- Birnbaum, M. H., & Mellers, B. A. (1988). Decision making and scaling studies. Working Paper. University of California.
- Birnbaum, M. H., Parducci, A., & Gifford, R. K. (1971). Contextual effects in information integration. *Journal of Experimental Psychology*, *88*, 158-170.
- Birnbaum, M. H., & Stegner, S. E. (1979). Source credibility in social judgment: Bias, expertise, and the judge's point of view. *Journal of Personality & Social Psychology*, *37*, 48-74.
- Birnbaum, M. H., & Stegner, S. E. (1981). Measuring the importance of cues in judgment for individuals: Subjective theories of IQ as a function of heredity and environment. *Journal of Experimental Social Psychology*, *17*, 159-182.
- Birnbaum, M. H., & Veit, C. T. (1974). Scale convergence as a criterion for rescaling: Information integration with difference, ratio, and averaging tasks. *Perception & Psychophysics*, *15*, 7-15.
- Breault, K. D. (1983). Psychophysical measurement and the validity of the modern economic approach: A presentation of methods and preliminary experiments. *Social Science Research*, *12*, 187-203.
- Coombs, C. H., Bezembinder, T. G., & Goode, F. M. (1967). Testing expectation theories of decision making without measuring utility or subjective probability. *Journal of Mathematical Psychology*, *4*, 72-103.
- Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, *51*, 380-417.
- Edwards, W., von Winterfeldt, D., & Moody, D. L. (1988). Simplicity in decision analysis: An example and a discussion. In D. Bell, H. Raiffa, & A. Tversky (Eds.), *Decision making: Descriptive, normative, and prescriptive interactions*. New York: Wiley.
- Ellsberg, D. (1961). Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics*, *75*, 643-649.
- Elmasian, R., & Birnbaum, M. H. (1984). A harmonious note on pitch: Scales of pitch derived from subtractive model of comparison agree with the musical scale. *Perception & Psychophysics*, *36*, 531-537.
- Fishburn, P. C. (1970). *Utility theory for decision making*. New York: Wiley.
- Fishburn, P. C. (1983). Research in decision theory: A personal perspective. *Mathematical Social Sciences*, *5*, 129-148.
- Galanter, E. (1962). The direct measurement of utility and subjective probability. *American Journal of Psychology*, *75*, 208-220.
- Galton, F. (1907). One vote, one value. *Nature*, *75*, 414.
- Goldstein, W., & Einhorn, H. J. (1987). A theory of preference reversals. *Psychological Review*, *94*, 236-242.
- Grether, D. M., & Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *American Economic Review*, *69*, 623-638.
- Hagerty, M., & Birnbaum, M. H. (1978). Nonmetric tests of ratio vs. subtractive theories of stimulus comparison. *Perception & Psychophysics*, *24*, 121-129.
- Hardin, C., & Birnbaum, M. H. (1990). The malleability of "ratios" of occupational prestige. *American Journal of Psychology*, *103*, 1-20.
- Hershey, J. C., Kunreuther, H. C., & Schoemaker, P. J. (1982). Sources of bias in assessment procedures for utility functions. *Management Science*, *28*, 936-954.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263-291.
- Karmarkar, U. S. (1978). Subjectively weighted utility: A descriptive extension of the expected utility model. *Organizational Behavior and Human Performance*, *21*, 61-72.

- Karni, E., & Safra, Z. (1987). "Preference reversal" and the observability of preferences by experimental methods. *Econometrica*, *55*, 675-685.
- Keeney, R., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs*, New York: Wiley.
- Keller, L. R. (1985). An empirical investigation of relative risk aversion. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-15*, 475-482.
- Knetsch, J. L., & Sinden, J. A. (1984). Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value. *Quarterly Journal of Economics*, *99*, 507-521.
- Komorita, S. S. (1964). A model for decision-making under risk. *American Journal of Psychology*, *77*, 429-436.
- Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement* (Vol. 1). New York: Academic Press.
- Kruskal, J. B., & Carmone, F. J. (1969). MONANOVA: A FORTRAN-IV program for monotone analysis of variance. *Behavioral Science*, *14*, 165-166.
- Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, *89*, 46-55.
- Luce, R. D. (1986). Comments on Plott and on Kahneman, Knetsch, and Thaler. *Journal of Business*, *59*, S337-S343.
- Luce, R. D., & Fishburn, P. C. (1991). Rank- and-sign dependent linear utility models for finite first order gambles. *Journal of Risk and Uncertainty*, *4*, 29-59.
- Luce, R. D., & Narens, L. (1985). Classification of concatenation measurement structures according to scale type. *Journal of Mathematical Psychology*, *29*, 1-72.
- Machina, M. J. (1982). Expected utility analysis without the independence axiom. *Econometrica*, *50*, 277-323.
- Mellers, B. A., & Birnbaum, M. H. (1982). Loci of contextual effects in judgment. *Journal of Experimental Psychology: Human Perception & Performance*, *8*, 582-601.
- Mellers, B. A., & Birnbaum, M. H. (1983). Contextual effects in social judgment. *Journal of Experimental Social Psychology*, *19*, 157-171.
- Mellers, B. A., Davis, D., & Birnbaum, M. H. (1984). Weight of evidence supports one operation for "ratios" and "differences" of heaviness. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 216-230.
- Mellers, B. A., Ordóñez, L., & Birnbaum, M. H. (1992a). A change of process theory of contextual effects and preference reversals in risky decision making. *Organizational Behavior and Human Decision Processes*.
- Mellers, B. A., Weiss, R., & Birnbaum, M. H. (1992b). Violations of dominance in pricing judgments. *Journal of Risk and Uncertainty*, *5*, 73-90.
- Miyamoto, J. M. (1983). An axiomatization of the ratio/difference representation. *Journal of Mathematical Psychology*, *27*, 439-455.
- Miyamoto, J. M. (1989). Generic utility theory: Measurement foundations and applications in multiattribute utility theory. *Journal of Mathematical Psychology*, *32*, 357-404.
- Narens, L., & Luce, R. D. (1986). Measurement: The theory of numerical assignments. *Psychological Bulletin*, *99*, 166-180.
- Parker, S., Schneider, B., Stein, D., Popper, R., Darte, E., & Needel, S. (1981). Utility function for money determined using conjoint measurement. *American Journal of Psychology*, *94*, 563-573.
- Payne, J. W. (1973). Approaches to decision making under risk: Moments versus risk dimensions. *Psychological Bulletin*, *80*, 439-463.
- Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization*, *3*, 324-345.
- Raiffa, H. (1968). *Decision analysis*. Reading, MA: Addison-Wesley.

- Rose, B. J., & Birnbaum, M. H. (1975). Judgments of difference and ratios of numerals. *Perception & Psychophysics*, *18*, 194–200.
- Rule, S. J., & Curtis, D. W. (1982). Levels of sensory and judgmental processing: Strategies for the evaluation of a model. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 107–122). Hillsdale, NJ: Erlbaum.
- Savage, L. J. (1954). *The foundations of statistics*. NY: Wiley.
- Schoemaker, P. J. (1982). The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature*, *20*, 529–563.
- Shanteau, J. (1974). Component processes in risky decision making. *Journal of Experimental Psychology*, *103*, 680–691.
- Shanteau, J. (1975). An information integration analysis of risky decision making. In M. Kaplan & S. Schwartz (Eds.), *Human judgment and decision processes* (pp. 109–137). New York: Academic Press.
- Slovic, P., & Lichtenstein, S. (1983). Preference reversals: A broader perspective. *American Economic Review*, *73*, 596–605.
- Slovic, P., Lichtenstein, S., & Fischhoff, B. (1988). Decision making. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology* (Vol. 2). New York: Wiley.
- Stevens, S. S. (1975). *Psychophysics*. New York: Wiley.
- Stigler, G. J. (1950a). The development of utility theory. I. *Journal of Political Economy*, *58*, 307–327.
- Stigler, G. J. (1950b). The development of utility theory II. *Journal of Political Economy*, *58*, 373–396.
- Tversky, A. (1967a). Additivity, utility, and subjective probability. *Journal of Mathematical Psychology*, *4*, 175–201.
- Tversky, A. (1967b). Utility theory and additivity analysis of risky choices. *Journal of Experimental Psychology*, *75*, 27–36.
- Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*, *59*, S251–S278.
- Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, *95*, 371–384.
- Veit, C. T. (1978). Ratio and subtractive processes in psychophysical judgment. *Journal of Experimental Psychology: General*, *107*, 81–107.
- von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton, NJ: Princeton Univ. Press.
- von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research*. Cambridge: Cambridge Univ. Press.
- Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, *55*, 95–115.

RECEIVED: November 6, 1989