Preference Reversals: A Broader Perspective

By Paul Slovic and Sarah Lichtenstein*

Two papers recently published in this Review, the first by Werner Pommerehne, Freidrich Schneider, and Peter Zweifel (1982) and the second by Robert Reilly (1982), re-examined the preference reversal phenomenon. Preference reversals occur when individuals are presented with two gambles, one featuring a high probability of winning a modest sum of money (the P bet), the other featuring a low probability of winning a large amount of money (the $ bet). The typical finding is that people often choose the P bet but assign a larger monetary value to the $ bet. This behavior is of interest because it violates almost all theories of preference, including expected utility theory.

The studies by Pommerehne et al. and Reilly were based on an earlier paper appearing in this Review by David Grether and Charles Plott (1979). All three of these investigations have followed the same general design, motivated by a healthy skepticism of the phenomenon and a belief that, examined under proper conditions, it might disappear. Thus Grether and Plott took great pains to correct what they saw as deficiencies in the original psychological experiments by ourselves (1971, 1973) and Harold Lindman (1971). Specifically, Grether and Plott used two monetary incentive systems to heighten motivation, substituted a different probability device for deciding the outcomes of the bets, controlled for income and order effects, and tested for indifference and the influence of strategic or bargaining effects. To their surprise, preference reversals remained much in evidence, despite their careful attempts to create conditions that would minimize or eliminate them.

Pommerehne et al., not satisfied with the stringency of Grether and Plott’s controls, attempted to increase motivation by raising the face value of the payoffs and creating differences in expected value between the P and $ bets in a pair. They, too, found a substantial proportion of reversals, leading them to conclude: “Even when the subjects are exposed to strong incentives for making motivated, rational decisions, the phenomenon of preference reversal does not vanish” (p. 573).

Reilly was also skeptical of the adequacy of Grether and Plott’s controls. To maximize subjects’ understanding of the task, he conducted his study within small groups where questions could readily be asked of the experimenter. The money at risk was placed on a desk in front of the subject and the size of potential losses in the gambles was increased to enhance motivation. Finally, some subjects were shown the expected values for all gambles and were given a description of the expected-value concept. Although the rate of preference reversals was somewhat lower than that observed by Grether and Plott, the phenomenon persisted to a substantial extent. Reilly conceded that these results provided “further confirmation of preference reversal as a persistent behavioral phenomenon in situations where economic theory is generally applied” (p. 582). Nevertheless, he maintained the hope that further strengthening of monetary incentives and provision of additional information to the subjects would make this troublesome phenomenon disappear, thus salvaging preference theory:

Should sufficiently large reductions be achievable, we might consider adopting the premise that individuals are likely to be consistent in making decisions that matter to them when the principle characteristics of the alternatives are sufficiently comprehended. Applied to such cases, standard preference theory would then require little modification.

*Decision Research, A Branch of Perceptrons, 1201 Oak Street, Eugene, Oregon 97401. The work was supported by the Office of Naval Research under Contract N00014-82-C-0643 to Perceptrons, Inc. We thank Don MacGregor, Amos Tversky, and two anonymous reviewers for comments on an earlier draft.
As researchers who have studied preference reversals and related problems of rational choice for quite some time, we have several concerns about the direction this research seems to be taking. Certainly a phenomenon such as preference reversals should be subjected to rigorous tests such as those administered by Grether and Plott, Pommerehne et al., and Reilly. These studies have been valuable in demonstrating the robustness of the effect. However, there is a substantial body of research on preference reversals within the psychological literature that is being neglected here. Moreover, reversals can be seen not as an isolated phenomenon, but as one of a broad class of findings that demonstrate violations of preference models due to the strong dependence of choice and preference upon information processing considerations. In this paper we shall describe relevant psychological work in order to broaden the perspective on preference reversals.

I. History

Readers of the papers by Pommerehne et al. and Reilly would hardly know there was considerable scrutiny of preference reversals prior to the publication by Grether and Plott. In fact, a number of studies preceded Grether and Plott, most of which employed multiple experiments and conditions designed to test the robustness of the effect. Additional studies have appeared subsequently. Each of these studies has observed substantial frequencies of reversals.

The first study designed to elicit reversals was our 1971 article. The impetus for this study was our observation in our earlier 1968 article that choices among pairs of gambles appeared to be influenced primarily by probabilities of winning and losing, whereas buying and selling prices were primarily determined by the dollar amounts that could be won or lost. When subjects found a bet attractive, their prices correlated predominantly with the amount to win; when they disliked a bet, their prices correlated primarily with the amount that could be lost. This pattern of correlations was explained as the result of a starting point (anchoring) and adjustment procedure used when setting prices. Subjects setting a price on an attractive gamble appeared to start with the amount to win and adjust it downward to take into account the probability of winning and losing, and the amount that could be lost. The adjustment process was relatively imprecise, leaving the price response greatly influenced by the starting point payoff. Choices, on the other hand, appeared to be governed by different rules.

In our 1971 article, we argued that, if the information in a gamble is processed differently when making choices and setting prices, it should be possible to construct pairs of gambles such that people would choose one member of the pair but set a higher price on the other. We proceeded to construct a small set of pairs that clearly demonstrated this predicted effect. Following this, a second study was conducted to examine the strength of the reversal effect as a function of the characteristics of the bet pairs. Forty-nine pairs of bets were constructed, all constrained by the requirement that the P bet had a high probability of winning a modest amount and the S bet had a low to moderate probability of winning a large amount. Despite these constraints, the pairs differed significantly in the degree to which they elicited predictable reversals. The ideal bet pair for observing reversals had a larger S bet loss than a P bet loss (facilitating choice of the P bet) and a large S bet win relative to the P bet win (facilitating a higher price for the S bet). For example, the bet with the most predicted reversals was: P bet, 9/12 to win $1.20 and 3/12 to lose $.10; S bet, 3/12 to win $9.20 and 9/12 to lose $2.00. We concluded this initial study by noting that reversals were of interest not only because they violated theories of rational choice, but because of the insight they revealed about the nature of judgment and decision processes.

Our 1968 article also noted that the close dependence of pricing responses on a gamble's payoffs could explain a finding that had

1Contrary to the explanation by Reilly, the act of choosing the P bet but setting a higher price on the S bet is not called a predicted reversal simply because “In all experiments reversal of P bets has been more frequent than reversal of $ bets…” (Reilly, 1982, p. 577, fn. 2).
puzzled Lindman (1965) in his doctoral dissertation. Lindman’s subjects gave selling prices for gambles and also made paired-comparison choices among triplets of these gambles. He noted that the prices were ordered almost perfectly according to the payoffs, whereas the orderings derived from choices were not. Lindman (1971) subsequently performed five studies designed to determine whether this sort of inconsistency would be influenced by the number of gambles within the choice set, the possibility of comparing gambles directly when deciding upon selling prices, variations in the way that probabilities were displayed, and variations in the amount of prior practice or experience. Although the experience factor had some effect, the general results across conditions were in close agreement with our own.

Problems of motivation and understandability were of concern right from the beginning of these studies. Experiment III of our original paper (1971) allowed college student subjects to win up to $8, a significant amount for an hour’s work in 1969. Each subject was run individually, with lengthy and careful instructions. Prices and choices were obtained three times for each pair of bets. The third time, subjects were reminded what their earlier answers had been and were asked to make a careful, final response. The bets were actually played and subjects were paid as a function of their winnings. Results for these carefully trained and financially motivated subjects showed a substantial proportion of predicted reversals. Recognizing the importance of motivation and the need to test nonstudent subjects, we went to considerable effort to replicate the initial studies on the floor of a casino in downtown Las Vegas.2 There the players could set the value of their chips at $0.05, $0.10, $0.25, $1, or $5. No players ever chose $1 or $5, but even for the $.10 chips, a typical $1 bet offered either a win or a loss of $8 on a single play. One new feature of the design was the addition of gambles having negative expected values. The experiment attracted 44 players, many of whom were highly educated professionals. Reversals of preference were frequent and widespread across players, even for the negative expected value gambles, for which strategic tendencies to overprice the bets would have worked against the reversal phenomenon.

Robert Hamm (1979) was another researcher who tried hard to make the reversal phenomena disappear—and did not. His extensive study examined the stability of reversals over time, in the face of experience, practice, forced introspection or discussion, and advice to adopt an intuitive or analytic approach to the task. The order of stimulus sets and tasks was carefully counterbalanced. Hamm found that the reversal effect was replicated under all these conditions. Task order had no effect, nor did emphasis on analytic or intuitive processes. Discussion about one’s decision strategies actually increased the tendency towards reversals, countering the hypothesis that if people were given greater opportunity to think about their strategies, the preference reversal phenomenon would disappear.

John Mowen and James Gentry (1980) studied preference reversals in a quite different context—that of new product development. Their subjects were undergraduate students of marketing and consumer behavior. They also extended previous research by comparing individual vs. group decisions. The stimuli were hypothetical products, defined according to probability of success and failure, and the projected profits and losses associated with those probabilities. Although the proportion of reversals varied with the characteristics of the pairs, as found in our (1971) study, strong reversal effects were generally observed. Group judgments and decisions were even more prone to reversals than those of individuals. Because group decisions involve discussion of strategies, this result is congruent with the effects of discussion found by Hamm. Mowen and Gentry related the anchoring process thought to determine pricing responses to an anecdote provided by R. A. Kerr (1979) regarding the search for oil in the Baltimore Canyon. Kerr noted that oil companies paid $1.1 billion for the privilege of drilling despite negative reports from oil industry geochemists. He con-

---

2 See our article (1973).
cluded that “Company managers apparently bid more on the basis of how large the possible trapping structures were rather than on the basis of the odds figured by the geochemists” (p. 1071).

In sum, many of the concerns raised and examined by Grether and Plott have also been investigated in other studies of preference reversals. Our purpose in reviewing these studies is not to deny the importance of the studies by Grether and Plott, Pomerenehne et al., and Reilly, but rather to inform those interested in this topic about the larger body of results. In our opinion, the most striking result of these studies is the persistence of preference reversals in the face of determined efforts to minimize or eliminate them.

II. A Broader View of Preference Reversals

The inconsistency between prices and choices for risky prospects represents but one of a broad set of failings that have been attributed to the theory of rational choice. James March (1978, 1982) has identified five general problems with the theory, one of which is particularly relevant to the present discussion. According to March, the theory presumes two improbably precise guesses about the future. One is a guess about the future consequences of current actions. The other is a guess about future sentiments (i.e., preferences) with respect to those consequences.

March (1978) argued that, partly as a result of behavioral research on human information-processing limitations, the way that the rational theory deals with the first guess has been modified to incorporate principles of what Herbert Simon (1957) termed “bounded rationality.” Thus economic theories now place considerable emphasis on notions of search, attention, and information costs. Aspiration levels, incrementalism, and satisficing have been described as sensible in many settings.

In contrast, March observed that although the second guess, about uncertain preferences, has so far had little effect in modifying normative theories, it poses potentially greater difficulties for these theories and their applications. He argued that limited cognitive capacity affects information processing about preferences just as it affects information processing about consequences: “Human beings have unstable, inconsistent, incompletely evoked, and imprecise goals at least in part because human abilities limit preference orderliness” (1978, p. 598).

March draws upon a rich and diverse array of observations to argue that, contrary to normative theory, preferences are neither absolute, stable, consistent, precise or exogenous (unaffected by the choices they control). The case against consistency brings us back to the topic of preference reversals. Inconsistencies between prices and choices were created on the basis of knowledge about different rules for processing the component aspects or dimensions of gambles. Since 1968, when information processing ideas began to be applied to risky choice, we have learned more about how perception and cognition determine preferences. As we have better understood those processes, it has become relatively easy, indeed almost commonplace, to produce new kinds of preference reversals. In many instances, production of reversals has been used to validate hypotheses about information processing in risky choice.

An early demonstration of the link between information processing and reversals was a study by Amos Tversky (1969). He hypothesized that, where the structure of the choice set permitted, it would be simpler and more natural to compare alternatives dimension by dimension than to evaluate the combined worth of each alternative separately (across dimensions) and then compare these overall evaluations. Tversky further hypothesized that small differences (for example, below some threshold of discrimination) would be ignored, even for an important dimension. Tversky tested and confirmed these hypotheses by creating sets of gambles in which this sort of information processing led to systematic, predictable intransitivities. Tversky’s gambles contained only two dimensions, probability of winning and amount to win. For his subjects, probability was the dominant dimension, but if the difference between gambles was small, then amount to win controlled the decision. Thus, given the
set of gambles \(a, b, c, d,\) and \(e\) with probabilities of \(7/24, 8/24, 9/24, 10/24,\) and \(11/24\) to win \$5.00, \$4.75, \$4.50, \$4.25,\) and \$4.00, respectively, \(a\) tended to be chosen over \(b, b\) over \(c, c\) over \(d,\) and \(d\) over \(e,\) presumably because the difference in payoffs outweighed the slight difference in probabilities within each of these pairs. However, \(e\) was typically chosen over \(a\) because of the relatively large difference in probabilities. This general finding has subsequently been replicated and extended by Rob Ranyard (1976) and by Lindman and James Lyons (1978).

The intransitivities observed by Tversky arose from the tendency of subjects to compare gambles on each dimension. If they had made holistic evaluations separately for each gamble and compared these to determine their choices, then the intransitivities would not have occurred. Comparison within dimensions is a natural way to choose among multidimensional objects. However, information is sometimes not available for each dimension, a situation that can lead to reversals. Consider, for example, the task of predicting which of two college students, \(A\) or \(B,\) would get the higher grade point average. Two test scores are available for each student, to serve as the basis for prediction. One score, English Skill, is available for both students. The other information is unique—Quantitative Ability for Student \(A\) and Need for Achievement for Student \(B\) as shown below (the means and standard deviations of each test are different but are known to the evaluator).

<table>
<thead>
<tr>
<th></th>
<th>Student (A)</th>
<th>Student (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need for Achievement</td>
<td>(-)</td>
<td>30</td>
</tr>
<tr>
<td>English Skills</td>
<td>90</td>
<td>131</td>
</tr>
<tr>
<td>Quantitative Ability</td>
<td>602</td>
<td>(-)</td>
</tr>
</tbody>
</table>

Slovic and Douglas MacPhillamy (1974) hypothesized that commonality would cause a dimension to be weighted more heavily in determining a choice, because common information is easier to use. This, in fact, occurred and led to systematic reversals on the above problem: 75 percent of the subjects rating the students individually gave a higher grade point average to Student \(A.\) However, when these same subjects were asked to make a comparative judgment, they selected Student \(B\) 60 percent of the time (40 percent of the subjects exhibited reversals). Reversals also occurred, though less frequently, when the means and standard deviations were the same for each test.

A variety of different reversals, providing strong evidence against traditional theories of preference, have come from the work of Daniel Kahneman and Tversky (1979; Tversky and Kahneman, 1981). From their systematic observations of choices among risky alternatives, Kahneman and Tversky have deduced a number of general principles, some of which violate expected utility theory, others of which are incompatible with all existing theories of choice or preference. Kahneman and Tversky distinguished between two phases in the choice process, an early phase of editing and a subsequent phase of evaluation. The editing phase, which they have also referred to as framing, consists of a preliminary analysis of the available options, their possible outcomes, and the contingencies or conditional probabilities relating outcomes to acts. One function of the framing process is to organize and reformulate the alternatives so as to simplify the second phase of evaluation and choice. Much as changes in vantage point induce alternative perspectives on a visual scene, the same decision problem can be subject to many alternative frames. Whichever frame is adopted is determined in part by the external formulation of the problem and in part by the standards, habits, and personal predilections of the decision maker.

A key element of framing is the coding of outcomes. Kahneman and Tversky show that, contrary to utility theory, outcomes are typically coded as gains and losses, rather than as final states of wealth. These gains and losses are defined relative to some neutral reference point, typically, but not always, the current asset position of the decision maker. These changes are evaluated according to a value function, \(v(x),\) which attaches a subjective worth to each possible outcome of a gamble, and a nonlinear probability weighting function, \(\pi(p),\) which expresses the subjective importance attached to the probability of obtaining a particular outcome. The attractiveness of a gamble that offers a chance
of \( p \) to obtain outcome \( x \) and a chance of \( q \) to obtain outcome \( y \) would be equal to \( \pi(p)v(x) + \pi(q)v(y) \). In addition to being defined on gains and losses relative to some psychologically meaningful (neutral) reference point, the value function is steeper for losses than for gains, meaning that a given change in one’s status hurts more as a loss than it pleases as a gain. Another important feature is that the function is concave above that reference point and convex below it, meaning, for example, that the subjective difference between gaining (or losing) $10 and $20 is greater than the difference between gaining (or losing) $110 and $120. Perhaps the most notable feature of the probability weighting function is the great importance attached to outcomes that will be received with certainty. Thus, for example, the prospect of losing $50 with probability of 1.0 is more than twice as aversive as the prospect of losing the same amount with probability .5.

The way a problem is framed determines both the reference point (the zero point) of the value function and the probabilities that are evaluated. If \( \pi \) and \( v \) were linear functions, preferences among options would be independent of the framing of acts, outcomes, or contingencies. Because of the characteristic nonlinearities of \( \pi \) and \( v \), however, normatively inconsequential changes in the frames significantly affect preferences. This is illustrated by the following pair of problems, given to separate groups of respondents.

**Problem 1.** Imagine that the United States is preparing for the outbreak of an unusual disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the consequences of the programs are as follows: If Program A is adopted, 200 people will be saved. If Program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved. Which of the two programs would you favor?

**Problem 2.** (Same cover story as Problem 1.) If Program C is adopted, 400 people will die. If Program D is adopted, there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die. Which of the two programs would you favor?

Although the two problems are formally identical, the preferences tend to be quite different. In a study of college students, 72 percent of the respondents chose Program A over Program B and 78 percent chose Program D over Program C. This difference can be traced to the different frames implied by the two problems. The “save lives” wording of the first problem implies that the value function’s reference point is the loss of 600 lives, while the “people will die” wording of problem 2 suggests that the reference point is at no lives lost. Thus problem 1 falls in the concave gain region of the value function while problem 2 is in the convex loss region.

Another study, surveying physicians and patients regarding choice of radiation vs. surgical treatments for lung cancer, produced different decisions when relevant statistics were changed from probabilities of surviving for various lengths of time after treatment to probabilities of not surviving (Barbara McNeil et al., 1982).

Another example of framing effects has been presented by Kahneman and Tversky (1982).

**Problem 1.** Imagine that, in addition to whatever else you have, you have been given $200. You are now asked to choose between (A) a sure gain of $50 and (B) a 25 percent chance of winning $200 and a 75 percent chance of winning nothing.

**Problem 2.** Imagine that, in addition to whatever you have, you have been given a cash gift of $400. You are now asked to choose between (C) a sure loss of $150, and (D) a 75 percent chance of losing $200 and a 25 percent chance of losing nothing.

Most people choose A over B and D over C. Yet, the options presented in the two problems are identical. There is no valid reason to prefer the gamble in one version and the sure outcome in the other. Choosing the sure gain in the first problem yields a total gain of $200 plus $50, or $250. Choosing the sure loss in the second version yields the same result through the deduction of $150 from $400. The choice of the gamble in either problem yields a 75 percent chance of winning $200 and a 25 percent chance of
winning $400. If the respondents to these problems took a comprehensive view of the consequences, as is assumed by theories of rational decision, they would combine the bonus with the available options and evaluate the composite outcome. Instead they ignore the bonus and evaluate the first problem as a choice between gains and the second as a choice between losses. The reversal of preferences is induced by reframing the problem.

We have used the framing and reference point notions to explain the finding that the certain loss of a stated amount of money (for example, $50) was much more attractive when described as an insurance premium (to safeguard against a .25 chance of losing $200) than when it was described as an alternative to playing that same gamble (see our article with Baruch Fischhoff, 1982a; see also Paul Schoemaker and Howard Kunreuther, 1979, and John Hershey and Schoemaker, 1980, for similar results).

III. Where Next?

We have presented a sample of the sorts of preference reversals that have formed our understanding of choice processes or have been created from that understanding. Those who are concerned about the possible economic implications of these phenomena have several paths to consider. One is to continue to subject these studies to the sorts of scrutiny that Grether and Plott and others have applied to the inconsistency between prices and choices. Despite the claims by Tversky and Kahneman (1981) that the effects they described are large and systematic, associated with losses of human life as well as monetary outcomes, not restricted to hypothetical questions, nor eliminated by monetary incentives, this line of research is young and there is certainly a need to test the limits and robustness of its findings.

A second path is to modify utility theory in order to accommodate as many of the behavioral anomalies as possible without abandoning the theory altogether. This has been a popular direction in recent years. A number of theorists have proposed weakening or eliminating the substitution axiom in order to accommodate the Allais paradox (Maurice Allais, 1953) and certain other violations of the traditional model (see S. H. Chew and Kenneth MacCrimmon, 1979; Peter Fishburn, 1981; Robert Weber, 1982; Hector Munera and Richard de Neufville, 1982; and Mark Machina, 1982). However, none of these revamped models can explain the framing effects described by Tversky and Kahneman (1981) or the preference reversals among $P$ bets and $S$ bets. Indeed, Machina acknowledged that, "to the extent that preference reversals are found to be systematic and pervasive, the behavioral model presented here must either be generalized or replaced" (p. 308).

A third path to follow, and one that we would advocate, is to accept the reality of preference reversals and related information-processing phenomena, and to explore their implications for important social and economic behaviors. We have begun to do this with regard to problems of societal risk management and programs for informing the public about risk (see our study, with Fischhoff, 1982b). Similarly, March (1978, 1982), whose critique went far beyond information processing to encompass complex strategic and social motivations, has urged that a conception of preference that respects the "intelligence of ambiguity" be incorporated into what he calls "the engineering of choice." He identified a number of conceptual problems that need to be addressed by choice theorists and optimization problems that need to be considered by choice engineers.

In a narrower but nonetheless important vein, Hershey, Kunreuther, and Schoemaker (1982) have demonstrated biases in utility functions caused by information processing effects. They showed that methods for assessing utilities, varying in normatively inconsequential ways, produced very different utility functions, posing both practical and theoretical problems for those concerned with assessing people's risk preferences. Donald Wehrung, MacCrimmon, and K. Brothers (1980) obtained similar inconsistencies with business executives, leading them to question the use of utility theory as a management tool. A more general analysis of the difficulties of assessing preferences has been presented by Fischhoff and ourselves (1980).
Fischhoff et al. argue that the strong effects of framing and information-processing considerations make elicitation methods major forces in shaping the expression of one's personal values.

Robin Gregory (1982) investigated a number of different approaches for estimating the value of nonmarket goods such as air and water quality, protection of threatened environments and species, and access to uninhabited views. He examined two measures of economic value; one based on an individual's willingness to pay to obtain or retain a good and the other based on the amount of compensation demanded if it is relinquished. He found that both methods were subject to sizable framing and information-processing effects.

Richard Thaler (1980) has drawn upon the reference point and framing notions of Kahneman and Tversky to explain a number of "economic illusions" that cause consumer behavior to deviate from the predictions of normative models. Included in his analysis were the overweighting of out-of-pocket costs relative to opportunity costs (foregone gains), the failure to ignore sunk costs, and the effects of psychic regret on such diverse areas as health care delivery decisions and vacation planning. Thomas Russell and Thaler (1982) argued that departures from rationality due to information-processing effects are unlikely to disappear in competitive markets. Kenneth Arrow (1982) underscored this argument by pointing out a number of failures of the rational model in insurance, securities, and futures markets that he feels are directly interpretable in terms of effects such as those linked to preference reversals and framing.

IV. Conclusion

This review has attempted to show how preference reversals fit into a larger picture of information-processing effects that, as a whole, pose a collective challenge to preference theories far exceeding that from reversals alone. These effects seem unlikely to disappear, even under rigorous scrutiny. Moreover, anything less than a radical modification of traditional theories is unlikely to accommodate these phenomena. We urge economists not to resist these developments but, instead, to examine them for insights into the ways that decisions are made and the ways that the practice of decision making can be improved.

REFERENCES


_____ and Schoemaker, Paul J. H., "Risk Taking and Problem Context in the Domain of Losses: An Expected Utility Anal-


, , and (1982b) *Regulation of Risk: A Psychological Perspective*, Decision Research Report 82–11, Deci-


