SECTION I. INTRODUCTION

1. The Construction of Preference: An Overview

Sarah Lichtenstein and Paul Slovic

This book is a collection of papers and research articles on *preference construction*. The central idea is that in many situations we do not really know what we prefer; we must construct our preferences as the situation arises.

We do, of course, carry a myriad of preferences in our memory. We were born with some of them, such as a fondness for sweets, avoidance of pain, and, perhaps, fear of snakes. Moreover, we spend our lives, particularly in childhood and adolescence, building preferences from our experiences and our wants, needs, and desires. Some of these learned preferences are broad, such as preferring more money to less; others are quite specific, such as liking a particular flavor of ice cream. These well-established preferences are readily available for use.

The need for preference construction arises when our known preferences are insufficient to solve the decision problem that we face. It seems to us that these more difficult situations have one or more of the following three characteristics.[1] First, some of the decision elements may be totally unfamiliar, such as when choosing what to eat from a menu in a foreign language. Second, the choices available to us may present a conflict among our known preferences, so that tradeoffs must be made. For example, suppose you are choosing between two apartments to rent. One has big windows with a great view but a cramped kitchen. The other has no view but a well-arranged, spacious kitchen. You know you prefer good views and large kitchens, but to make this decision you must make a tradeoff between one aspect (view) and the other (kitchens). Often we know our preferences for individual aspects but do not know the tradeoffs between them. Third, we find it difficult to translate our positive and negative feelings into a numerical response, even when our preferences are clear. For example, suppose you find an apartment with a great view and a big kitchen. Now the question is: How much rent are you willing to pay for this apartment? That question may be hard to answer.

Decision researchers rarely study the first problem, unfamiliarity. Indeed, researchers typically take great pains to choose stimuli that are familiar to the study participants. Thus, the emphasis in research on preference construction

[1] We thank John Payne for stimulating our thoughts on this.

is on the second problem, tradeoffs, or the third problem, numerical responses, or both. The distinction between these two problems is not always clear. For example, when you find the apartment with the great view and big kitchen, you might reframe the question as: How much more rent are you willing to pay for it than for the apartment with the small kitchen? This involves a tradeoff between money and kitchen space. But the distinction is valuable because one problem focuses on our internal processes in juggling our preferences and the other focuses on the interface between our preferences and our responses.

Decision research would be important but less interesting if we always used the same methods for constructing our preferences. But we do not. This book documents a huge variety of methods, strategies, and ways of thinking or feeling used to construct preferences. Moreover, both the choice of methods and the preferences themselves are determined not only by our knowledge, feelings, and memory but also by many aspects of the decision environment, including how the preference objects are described, how the preference question is posed, and what response is required.

The variability in the ways we construct and reconstruct our preferences yields preferences that are labile, inconsistent, subject to factors we are unaware of, and not always in our own best interests. Indeed, so pervasive is this lability that the very notion of a "true" preference must, in many situations, be rejected.

This book presents much evidence for this view of preference construction. In addition, it addresses a broad range of questions arising from the preference-construction thesis:

- How do we construct preferences? What techniques, mental juggling, or shortcuts do we use?
- What factors, internal and external, influence our preferences? How do these factors affect our choice of construction methods?
- Under what conditions are our construction methods good or bad (i.e., enhancing or detracting from our best interests)?
- How can our construction methods be improved?
- How should the evidence of preference construction be incorporated into theories of decision making?
- What are the personal, ethical, legal, and public-policy implications of the idea that "true" preferences do not exist?

Overview of This Chapter

The organization of this chapter is based on the contents of the other 37 chapters in this book, with a few side excursions. We start with *preference reversals*, that is, the documentation of situations in which, under one circumstance, A is preferred to B, whereas in another, seemingly equivalent, circumstance, B is preferred to A. The preference reversal literature was not the only, nor even

the first, precursor to the idea of preference construction, but it strongly influenced our own thinking and generated an enormous literature that is central to preference construction. In addition, it embodies both the difficulties that, as we suggested earlier, are the heart of preference-construction problems. In one circumstance, subjects are presented with a choice between two options, and they must resolve conflicts by making tradeoffs among the aspects in order to choose which option they prefer. In the other circumstance, subjects are presented with just one option (and then, later, the other), and their task is to translate their preference into a numerical response.

The next section of this chapter is a side excursion: the reaction by experimental economists to the findings of preference reversals. Although there is a substantial literature in economics on this topic, we have included only one such paper (Grether & Plott, 1979/Ch. 5) in this book. However, we believe we have a good perspective on the economic literature and so have indulged ourselves by presenting our views in this introductory chapter.

Following that is a brief section, "The Blossoming of Preference Reversals," that is also not paralleled by book chapters. This short section gives a sampling of the variety of stimuli and responses that have been used to show preference reversals. This is just a listing, without theoretical discussion, but its implication should be clear: The situations are so diverse that no single explanation will cover all of them, nor can any one experiment showing the elimination of preference reversals in one situation be taken as refutation of them all.

We then turn to the theories that have been developed to explain preference reversals, both the six theories included in this book (Chs. 6–11) and a few other theories developed around the same time. We end this section with a re-examination of explanations for the original preference reversal finding (i.e., Lichtenstein & Slovic, 1971/Ch. 3).

The next section of this chapter briefly describes the other precursors of the idea of preference construction, including the enormous influence of Herbert Simon (1955, 1956) on the early research on information processing and choice rules, leading to the first (to our knowledge) explicit statement of preference construction by Bettman (1979). Following that is a section describing the five studies (Chs. 12–16) we chose for this book to exemplify the variety of research that has explored aspects of preference construction.

We turn then to theories of preference construction. Five theories are presented here (Chs. 17–21). Because preference construction is now viewed as the core problem in decision making, virtually every current theory of decision making can be considered a theory of preference construction. Additional theories are presented in the next section (Chs. 22–26), which explores the interplay between the use of reason or analysis and the use of affect or emotion in decision making.

The next sections of this chapter and of the book are devoted to three special topics. The first illuminates discrepancies that Gilbert and Wilson (2000/Ch. 30) call miswanting: What we value now may not be reflected in our decisions, and

neither our current values nor our decisions may predict what we will value when we later experience the outcomes.

The second topic is a highly applied one: the use of people's reported willingness to pay (e.g., through higher taxes) to measure the economic value to our society of goods or programs that have no markets (e.g., the value of remote wilderness sites). From the perspective of preference construction, these methods, collectively called Contingent Valuation, have deep flaws that are explored here (Chs. 31–33).

The third topic is preference management. Given that our preferences are often labile, how might we manage them for our own good? An additional issue is that other people can influence our preferences by the ways in which the options are presented to us; in many situations this influence is unavoidable. What are the practical and ethical implications of these influences?

THE START OF PREFERENCE REVERSALS

When Paul Slovic, with his new Ph.D., moved from Ann Arbor, Michigan, to Eugene, Oregon, in 1964, he left Ward Edwards' world of decision making, gambles, and Subjectively Expected Utility to enter Oregon Research Institute's (ORI's) milieu, where Paul Hoffman, Lew Goldberg, and Len Rorer were using linear regression techniques to develop quantitative models of cue utilization in judgment (Hoffman, 1960). Combining these influences, Paul started to study how people weighted probabilities and payoffs when evaluating gambles, using the regression techniques pioneered at ORI (and also by Ken Hammond at Colorado; see Hammond, Hursch, & Todd, 1964). To obtain separate utilization weights for probabilities of winning and probabilities of losing, Paul devised the *duplex gamble*, which was depicted by two discs; one showed the probability of winning and the amount to win (otherwise win nothing); the other, the probability of losing and amount to lose (otherwise lose nothing). The gamble was played by spinning the pointers on both discs; the player could win and not lose, lose and not win, win and lose, or neither win nor lose (see Figure 1.1). Across a set of duplex gambles, the experimenter could independently vary the four risk dimensions of the bets: the (stated) amounts to win and lose and the (stated) probabilities of winning and losing. Thus, correlational techniques could be used to analyze the dimension weights underlying people's preferences for playing these gambles.

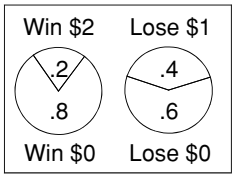

Figure 1.1. Example of a duplex bet.

Sarah Lichtenstein, who joined Paul at ORI in 1966, came from the same background and met the same change in paradigm. Our first major collaboration (Slovic & Lichtenstein, 1968; Ch. 2 is an abridgement) showed the melding of the decision making and judgment approaches: We used duplex bets and linear

regression models to examine whether two different response modes, bids (i.e., hypothetical buying and selling prices) and attractiveness ratings, led to systematic changes in the relative importance of the cues (i.e., the four risk dimensions).[2]

We developed a simple two-stage process model: First, decide whether the gamble is attractive or unattractive (the expected values of the gambles ranged from +$3.00 to −$3.00); second, provide a quantitative response (i.e., a rating or a price). For the first stage, we found that the probability of winning received the greatest weight. In this stage, the percentage of variance accounted for by each risk dimension was remarkably similar for rating and for pricing. But there were large response-mode differences in the second stage. For attractive bets, ratings were most influenced by the probability of winning, whereas prices were most influenced by the amount to win. For unattractive bets, no one risk dimension dominated in the rating data, but for prices the amount to lose captured 73% of the total variance accounted for by the linear model. We concluded that the regression weights may reflect beliefs about the relative importance of the risk dimensions and that "change in weights across tasks indicates changing strategies due to subjects' attempts to ease their information-processing problems" (this volume, p. 50).

We then reasoned that, because of the differing information-processing strategies apparently induced by different response modes, we could, with suitably chosen bets, create preference reversals. That is, we could construct pairs of gambles, A and B, such that A would be preferred under one response mode and B under another.

For our first test of this hypothesis, we selected pricing and choice as our two response modes. Prices for attractive bets, we knew, were based on the amount to win. We believed that choices would be based on attractiveness, so that, like the first stage of our 1968 model and like attractiveness ratings, choices would be more influenced by the probability of winning.[3]

Instead of using duplex gambles, which were unfamiliar and thus possibly suspect, we used ordinary two-outcome bets (with the probabilities of winning and losing summing to 1.00). For each pair, one bet, called the $ bet, featured a large amount to win, fostering a large bid, and the other bet, called the P bet, featured a high probability of winning, fostering its choice over the $ bet.

These were the origins of our first preference reversal paper (Lichtenstein & Slovic, 1971/Ch. 3), presenting three experiments. The first and third experiments showed very high rates of the kind of preference reversal we had intended

[2] In this 1968 work we *did not* collect any data about the relative importance of gamble attributes in choice. Many subsequent writings (starting with our own, Lichtenstein & Slovic, 1971) are wrong in reporting that we did. We apologize for this error.

[3] Later research has shown that the relationship between choice and rating responses is not this simple. For example, Goldstein and Einhorn (1987) found substantial discrepancies between these two response modes.

to create and very low rates of the opposite, unpredicted reversal (i.e., choosing the $ bet but bidding more for the P bet). For the third experiment, each subject was run individually, the bets were played, and subjects' pay was based on their winnings. At the end of the session, the experimenter interviewed the subjects who had made predicted reversals, trying to get them to change their responses to eliminate the reversals. A previously unpublished excerpt from one such interview is included in this book as an Appendix to Chapter 3.

The second experiment used more, and more varied, pairs of bets to explore the effect of changes in bet attributes on the frequency of preference reversals. We found that the difference between the amounts to lose in the $ bet and the P bet correlated .82 with the number of P-bet choices but was not related to bids (buying prices, in this case), whereas the ratio of the amounts to win in each pair was related ($r = .55$) to bids but not choices. The ranges of probability differences were too small to show any effects.

Our reasoning and results during those early years seemed to us to explain a finding that had puzzled Harold Lindman (1965). His subjects gave bids for a number of gambles and also made paired-comparison choices. He noted a consistent discrepancy between the preference orderings. After discussing these data with us, Lindman (1971) did a more systematic study comparing selling prices and choices and reached much the same conclusions as we had.

Soon after, we were given, through the generosity of Ward Edwards, an exceptional opportunity to replicate our findings in a downtown Las Vegas casino, the Four Queens (Figure 1.2). The experiment was run by John Ponticello, a dealer and pit boss. The subjects were casino patrons risking their own money. In this setting we were able to include bets with negative expected values; we used the mirror images of the positive bets. Thus, for example, the $ bet {2/12 to win 97 chips, 10/12 to lose 11 chips} became {10/12 to win 11 chips, 2/12 to lose 97 chips}. In accord with our 1968 data, we predicted that for bad bets subjects would choose the $ bet (indicating a lesser dislike for it) but offer a larger payment to the dealer to avoid playing the $ bet (indicating a greater dislike for it). Our predictions were strongly confirmed (Lichtenstein & Slovic, 1973/Ch. 4), replicating and extending the earlier findings based on college students.

ECONOMISTS' REACTIONS TO PREFERENCE REVERSALS

The existence of preference reversals presents a formidable challenge to economic theory because a fundamental tenet of economics is that prices and preferences are synonymous. If we prefer A to B, we should be willing to pay more for A. Accordingly, economists David Grether and Charles Plott (1979/Ch. 5) set out not merely to refute our findings, but "to discredit the psychologists' work as applied to economics" (this volume, p. 77). They listed 13 theories or explanations, most of which would, if confirmed, characterize preference

Figure 1.2. The editors in the Four Queens Casino in Las Vegas, Nevada, 1969.

reversals as unimportant artifacts. These included misspecified incentives (no real payoffs) and the fact that the previous researchers were psychologists (who "have the reputation for deceiving subjects," this volume, p. 85). They ran two experiments carefully designed to test all 13 possibilities. Finding a substantial frequency of predicted reversals (e.g., 70% for Experiment I with incentives) and far fewer unpredicted reversals (e.g., 13% in same group), they not only replicated our results but also rejected 12 of their 13 explanations. The data supported only our explanation based on information, processing changes due to changes in response mode.

Grether and Plott (1979/Ch. 5) issued a call-to-arms to economists:

> Taken at face value the data are simply inconsistent with preference theory . . . The inconsistency is deeper than the mere lack of transitivity or even stochastic transitivity. It suggests that no optimization principles of any sort lie behind even the simplest of human choices. (this volume, p. 77)

Economists and others responded to this alarm with vigor.[4] One approach was to modify utility theory. Loomes and Sugden (1983) and Fishburn (1985)

[4] According to the *Web of Science*, Grether and Plott's article had been cited in about 380 journal articles through February 2005. Our first preference-reversal article (Lichtenstein & Slovic, 1971/Ch. 3) had also been cited about 380 times; one or the other, or both, had been cited in about 600 articles.

explored utility theories without the transitivity assumption, attributing preference reversals to regret. Others focused on the practices, in many experiments, of having subjects actually play only one gamble, chosen at random, and using the Becker, DeGroot, and Marschak (1964) method for playing gambles. These practices can be seen as creating complicated, multistage lotteries. Segal (1988) suggested giving up the reduction principle (i.e., indifference between a multistage lottery and its single-stage equivalent). Holt (1986) and Karni and Safra (1987) proposed, instead, to explain preference reversals by retaining the reduction principle but giving up the independence axiom (i.e., only the outcomes that distinguish two lotteries are relevant to the decision – the axiom that is belied by the Allais, 1953, paradox). None of these approaches has been successful in explaining preference reversals (see Tversky, Slovic, & Kahneman, 1990/Ch. 8).

Philosophers have scolded economists for their approach to preference reversals, "whose existence was until quite recently denied by the majority of economists. Their favorite strategy consisted in trying to explain the phenomenon away as an artifact of the experimental techniques used to observe it" (Guala, 2000, p. 48). Guala tartly noted that "some of the best experimenters in economics devoted lots of time and effort to test the functioning of the [Becker–DeGroot–Marshak] and [Random Lottery Selection] mechanisms – despite the fact that . . . PR [preference reversals] had been observed *with* and *without* these elicitation procedures" (p. 50; emphasis in original). In a similar vein, Hausman (1991) titled his essay, "On Dogmatism in Economics: The Case of Preference Reversals." Hausman traced economists' reactions to their reluctance to abandon a broad and parsimonious theory of rational choice for psychologists' narrower and more complex theories. He concluded that economists' reactions were hard to defend, creating "unreasonable barriers to theoretical and empirical progress" (p. 223).

Changing the Task

One way to "explain away" preference reversals that has been used by many experimenters is to change the task, that is, to give subjects some tasks other than pricing bets or choosing between them or to change the rules of the game. We object to this approach because even if preference reversals were eliminated using the new task or the new rules, that finding would not refute preference reversals in the original tasks.

Moreover, there is a pattern to the results of these new-task studies: Preference reversals are not eliminated. Instead, predicted reversals decrease and unpredicted reversals increase. Such a finding is a comfort to economists because it is the asymmetry between high rates of predicted reversals and low rates of unpredicted reversals that challenges utility theory; symmetric reversal rates can be interpreted as utility maximization with error. However, the error explanation cuts both ways. We propose an alternative hypothesis: Suppose

that, if people never made errors, they would *always* make predicted reversals and *never* make unpredicted reversals. Then, as random errors increase, the rate of predicted reversals will be reduced and the rate of unpredicted reversals will increase. At one extreme is a task with simple, no-loss gambles (e.g., .5 to win $6.50; otherwise nothing) and easy response modes like attractiveness ratings and selling prices (without any selling mechanism specified). For this task, Goldstein and Einhorn (1987) found 93% reversals when the P bet was rated more attractive and only 2% reversals when the $ bet was rated more attractive. At the other extreme, our reversals-with-error hypothesis predicts that complex or confusing tasks will show a more symmetric pattern of reversals, as seen in the following five studies:

Cox and Epstein (1989) converted the pricing task to a kind of choice task. In addition, because of a peculiarity in their pseudo-pricing task, the bets used for pricing were different from the bets used in the plain choice task: For pricing, outcomes were lowered by 1,000 units. Thus, for example, the bet {.50 to win 6,500, .50 to lose 1,000} in choice became {.50 to win 5,500, .50 to lose 2,000}. The authors found equal rates of predicted and unpredicted reversals. In their first experiment, 34% of plain choices were inconsistent with "pricing" choices. Such a high rate of inconsistency between two forms of choice suggests that subjects may have been confused as to whether their pricing responses should reflect their choices (as the game dictated) or should be real prices.

Bostic, Herrnstein, and Luce (1990), with computer presentation and intermixed trials, used three response modes: choice, selling prices, and choice indifference points (i.e., choice between a bet and a sure outcome; the sure outcome is varied between trials until indifference is found). In their second experiment, the narrowing-in on the choice indifference point was well concealed from the subjects (cf. Fischer, Carmon, Ariely, & Zauberman, 1999, who made the narrowing-in obvious, with quite different results). Bostic et al. found the usual preference reversal results for prices versus choices, 81% predicted and 3% unpredicted conditional reversals, but the hidden narrowing-in process apparently increased errors; there were 51% predicted and 22% unpredicted reversals between choices and this kind of certainty equivalence.

Wedell and Bockenholt (1990), using choice and selling price, intermixed trials in which the subject was instructed to imagine that the bet (or the chosen bet) would be played once, 10 times, or 100 times. Again, the results support our reversals-with-error hypothesis; predicted and unpredicted reversal rates were approximately equal.

Cox and Grether (1996) looked at choices, selling prices, second-price sealed bid auctions (for five rounds), and English clock auctions (for five rounds). They found the usual preference reversals for prices and choices (61% predicted, 10% unpredicted) and, for the auctions, decreasing rates of predicted reversals over the five rounds, with increasing (sealed bid) or always high (English clock) rates

of unpredicted reversals. The English clock auction was apparently especially difficult for the subjects (the computer lowered the price every 5 seconds): In the first round, two-thirds of all responses were inconsistent with choices; in the fifth round, 45% were. This study is widely cited as showing that market experience reduces preference reversals. We think it shows that these markets induced more errors.[5]

Berg, Dickhaut, and Rietz (2003) used an extremely complex two-stage game and concluded that their data supported the hypothesis that subjects maximized utility with errors. We agree about the errors but believe that the data support our reversals-with-error explanation, not a utility-theory explanation.

We know of only one task that really does eliminate preference reversals. Chu and Chu (1990) required their subjects to engage in arbitrage (the money-pump game). In their second experiment, subjects, run individually, were presented with a single pair of bets; they made a choice and stated a price for each bet. If their responses showed a reversal, they immediately played the money-pump game with the experimenter for one round.[6] They then gave a new choice and new prices to the same bet pair and, if a preference reversal was again shown, they again played the money-pump game for one round. This cycle was continued until the subject did not reverse. The whole process was repeated for two more bet pairs. The subjects learned very quickly, requiring, on average, only 1.8 arbitrage experiences to eliminate their first predicted reversal and only 1.2 arbitrage experiences to eliminate their first unpredicted reversal (if any). Moreover, only 1 of the 40 persons who reversed on the first or second bet pair, and had thus experienced at least one arbitrage, reversed on the third bet pair.

Cherry, Crocker, and Shogren (2003) also found large reductions in rates of reversals (down to about 5%; they did not report predicted and unpredicted reversals separately) using arbitrage once per reversal over 10 pairs of bets. Moreover, they concurrently presented 10 other pairs of bets not subject to arbitrage; with these bets, reversal rates were reduced to about 10%.

These arbitrage results indicate that individuals can learn not to overvalue $ bets if they are repeatedly punished for doing so. Whether these data serve to restore the idea of a "true preference" is open to question. Braga and Starmer

[5] The effect of markets on preference reversals may be more complicated than this. Braga and Starmer (2004) compared second-price auctions with second-to-last price auctions. They found that between the first and fifth rounds the second-price auctions greatly reduced predicted reversals and greatly increased unpredicted reversals but the second-to-last priced auctions showed only moderate (e.g., from 82% to 60% in Experiment II) decreases in predicted reversals and no significant increases in unpredicted reversals.

[6] The experimenter (E) owns both bets. E sells the higher-priced bet to the subject (S) for S's stated price. Then E trades bets with S. (This trade is accepted by S because S has indicated a choice-preference for the bet E still owns over the bet S just bought.) Then E buys the lower-priced bet from S for S's stated price. At the end of this sequence, E again owns both bets but S is poorer by the difference between the two prices, having bought the higher-priced bet and sold the lower-priced bet.