MANAGEMENT SCIENCE
Vol. 35, No. 4, April 1989
*Printed in U.S.A.*

# BIAS IN UTILITY ASSESSMENTS: FURTHER EVIDENCE AND EXPLANATIONS*

ERIC J. JOHNSON AND DAVID A. SCHKADE
*The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania* 19104
*College of Business Administration, University of Texas, Austin, Texas* 78712

Judgments about simple gambles, such as those used in utility assessment, can generate sizable and systematic bias. After Hershey and Schoemaker (1985), we employ both the probability and certainty equivalence methods to explore bias. Our results show that: (1) the direction and degree of bias depend on characteristics of the assessment gamble such as the reference probability and the difference between outcomes, (2) presenting subjects with explicit anchors can change the size and direction of the bias, and (3) subjects using heuristic response strategies show significantly more bias than those using expectation strategies. We also discuss the status of possible explanations for the bias, in light of these new results, including PE mode reframing, random error, and anchoring and adjustment.
(DECISION ANALYSIS; UTILITY ASSESSMENT; DECISION-MAKING; HEURISTICS)

## Introduction

Many approaches to aiding or modeling decisions attempt to measure a decision maker's value or utility for choice alternatives. By asking several seemingly simple questions, an analyst constructs a representation of a decision maker's preferences. Because these decomposed judgments are less demanding than the complete decision problem, decision analytic models are thought to be superior to unaided holistic choice.

However, there is growing evidence that even these simple judgments might contain systematic error. Responses to these questions do not resemble a simple retrieval of the utility placed on a level of an attribute. Instead, responses seem to be the result of a *constructive* process (Bettman 1979) in which characteristics of the decision problem and decision strategy partially determine the response. For example, there is a growing literature showing that revealed preferences are affected by the response mode (Slovic and Lichtenstein 1983; Hershey, Kunreuther and Schoemaker 1982), and the probability used in elicitation (Karmarkar 1978; McCord and de Neufville 1984). Because these are normatively irrelevant features of the elicitation task, such results are a failure of *procedure invariance* (Tversky, Sattath and Slovic 1988).

Understanding the origins of these systematic errors is an important goal for prescriptive and descriptive decision research. For prescription, such research could help in developing better elicitation methods, and provide insights into decision errors that might occur in other tasks. For description, such research could provide linkages to other response mode effects such as preference reversals (Tversky, Sattath and Slovic 1988; Goldstein and Einhorn 1987; Lichtenstein and Slovic 1971, 1973). In this paper we examine bias using a procedure developed by Hershey and Schoemaker (1985) to explore the origins of inconsistencies in responses to utility assessment questions.

In utility assessment, a decision maker is often presented with the simplest of all risky choices, involving a sure outcome and a single two-outcome lottery. These problems consist of

- A certain or "sure" outcome, $S$,
- An uncertain outcome of greater value, $G$, where $G > S$,

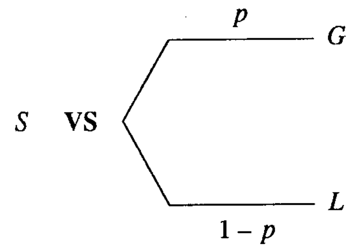
* Accepted by Robert L. Winkler, former Departmental Editor; received January 14, 1986. This paper has been with the authors 18 months for 3 revisions.



0025-1909/89/3504/0406$01.25

- An uncertain outcome of lesser value, $L$, where $L < S$,
- A probability, $p$, of obtaining $G$, and $q$, the probability of obtaining $L$, equal to $1 - p$.

Note that $S$, $G$, and $L$ can be either gains or losses, but that $G > S > L$.




A series of questions are asked in which three of the four elements ($S$, $G$, $L$, $p$) are specified by the analyst, and the decision maker must specify a value of the missing element that creates indifference between the risky and certain alternatives. For example, Probability Equivalence (PE) assessments ask the decision maker to set $p$ to a value creating indifference, for a specified $S$, $G$, and $L$. Another normatively equivalent method is Certainty Equivalence (CE), which asks the decision maker to report a value of $S$ that creates indifference, for a specified $p$, $G$, and $L$.

In Hershey and Schoemaker's procedure, decision makers answer assessment items in both the PE and CE response modes, with the answer from the first question serving as a gamble element in the second. For example, suppose that decision makers first complete a CE question that asks if they are indifferent between the gamble and its expected value. If not, they then respond with a different value for $S$. For an item with $G = \$100$, $L = \$0$, and $p = 0.5$, decision makers are asked if they are indifferent between the gamble and a sure \$50. A risk averse decision maker would answer no and respond with a value less than \$50 (say \$40). In a subsequent session, a PE question is asked with the same values of $G$ and $L$, but with \$40 for $S$, with $p$ to be specified. In our example, an item with $G = \$100$, $L = \$0$, and $S = \$40$ would be presented, and the decision maker would be asked if a probability of winning of 0.4 (the risk neutral probability equivalent) would produce indifference. To be consistent, the decision maker's PE should equal the initial probability ($p = 0.5$) presented in the first assessment. Any systematic discrepancy between the original $p$ and second-stage PEs represents bias. According to the axioms of expected utility theory, and more general notions of consistency, this second response should regenerate the original CE gamble.

Hershey and Schoemaker's results are marked by two systematic patterns of bias that occur in both orderings of response modes (CE-PE and PE-CE). First, second-stage responses are too high, in the utility sense. For gains, there is an overall positive bias: second-stage CE responses are higher than the initial sure amount $S$, and second-stage PE responses are higher than the initial probability, $p$. For losses, the pattern is similar: CE responses are more positive than $S$, and PE values are again higher than $p$.[1] Thus, gambles generated by second-stage responses dominate, on average, those originally completed in the first stage.

This overall positive bias is moderated by a second systematic pattern: an association between risk attitude, as measured by the first-stage response, and the presence of bias. Significant bias occurs only for certain combinations of risk attitude and response mode order producing an interaction: For the PE-CE case only risk averse subjects show bias; for CE-PE, only risk seeking subjects show bias.

[1] Hershey and Schoemaker used "one-outcome" gambles: for gains $L = 0$, and $G > 0$; for losses, $G = 0$, and $L < 0$. Thus, in the loss domain, decision makers were asked to specify $q$, the probability of $L$. These PE responses were consistently lower than $1 - p$, thus $p$ would be too high.

These combined results, an overall positive bias and an association with first stage risk attitude constrain possible explanations for this effect. Hershey and Schoemaker examine several possibilities, including PE mode reframing, anchoring and adjustment, and random error. While several explanations can produce the first result, an overall positive bias, Hershey and Schoemaker argue that only the PE reframing model also explains the observed interactive relationship between risk attitude and bias. According to this account, the sure amount, $S$, serves as a reference point in PE, and the two outcomes are recoded as a gain of $G - S$ and a loss of $S - L$.

In this paper we report three experiments that use Hershey and Schoemaker's procedure to further examine the causes and generality of this effect. These experiments extend the range of items and vary details of the elicitation process. Because the results of these experiments modify the description of the bias, we will examine a different theoretical account based on self-generated anchors. To explore the processes underlying bias, we will also collect process-tracing data using verbal reports and by monitoring information acquisition. In conclusion, we will examine the status of possible explanations of the bias in light of these new results.

## Experiment 1

### *Overview*

Experiment 1 has two primary goals: (1) extend the range of stimuli beyond that used by Hershey and Schoemaker (1985), by using probabilities other than 0.5 and gambles with two nonzero outcomes, and (2) extend their results to the case where explicit anchors are not provided.

The reference probability used in elicitation can affect revealed preferences (McCord and de Neufville 1984; Karmarkar 1978). For example, McCord and de Neufville found that higher reference probabilities generated more risk averse utility functions than lower probabilities. In this study we expand the set of first-stage reference probabilities used in elicitation from the single value used by Hershey and Schoemaker (0.5), to three levels (0.3, 0.5, 0.7). For PE, we manipulate an equivalent feature of the gambles, the value of $S$, relative to $G$ and $L$ (Table 1).

Also, we look at more complex items to address two concerns. First, we are interested in the generality of the effect, since many assessment procedures use two or more nonzero outcomes, while Hershey and Schoemaker's stimuli were gambles with one nonzero outcome. Second, increasing the complexity of the task may increase the use of simplifying heuristic strategies (Payne 1976) and this could, in turn, lead to an increase in decision errors such as bias (Johnson, Payne, and Bettman 1988). This change requires that we unconfound the difference between the two outcomes and the expected value of the gamble, factors which are, by definition, confounded when one outcome is zero, and $p$ is always 0.5. Using two nonzero outcomes therefore introduces another experimental factor, the difference between $G$ and $L$, which we term *Spread*.

We also depart from Hershey and Schoemaker's practice of asking subjects if the risk neutral response made them indifferent, since many utility assessment techniques do not use such a procedure (Farquhar 1984). Finally, like Hershey and Schoemaker, we examine both gains and losses, and both response mode orders: CE-PE and PE-CE.

### *Method*

The items were constructed using a factorial design that manipulated Domain (gains vs. losses), Spread ($G - L$, at 3 levels nested in each Domain), and Initial Probability ($p$, at 3 levels: 0.3, 0.5, and 0.7) (Table 1). For PE questions in the PE-CE condition, $p$ is not shown to subjects, but is implied by the value of $S$, relative to $G$ and $L$. Subjects completed the nine gain items, read similar instructions for losses, and then completed

TABLE 1

*Stimuli for Experiment* 1

| Gamble | $S$ | $G$ | $L$ | $p$ | Spread* |
|---|---|---|---|---|---|
| Gains | | | | | |
| 1 | 300 | 550 | 190 | 0.30 | 360 |
| 2 | 300 | 480 | 120 | 0.50 | 360 |
| 3 | 300 | 410 | 45 | 0.70 | 360 |
| 4 | 200 | 330 | 145 | 0.30 | 190 |
| 5 | 200 | 295 | 105 | 0.50 | 190 |
| 6 | 200 | 255 | 70 | 0.70 | 190 |
| 7 | 100 | 160 | 75 | 0.30 | 90 |
| 8 | 100 | 145 | 55 | 0.50 | 90 |
| 9 | 100 | 125 | 40 | 0.70 | 90 |
| Losses | | | | | |
| 10 | −350 | −200 | −415 | 0.30 | 215 |
| 11 | −350 | −245 | −455 | 0.50 | 215 |
| 12 | −350 | −285 | −500 | 0.70 | 215 |
| 13 | −250 | −50 | −335 | 0.30 | 285 |
| 14 | −250 | −110 | −390 | 0.50 | 285 |
| 15 | −250 | −165 | −450 | 0.70 | 285 |
| 16 | −150 | −65 | −185 | 0.30 | 115 |
| 17 | −150 | −90 | −210 | 0.50 | 115 |
| 18 | −150 | −110 | −240 | 0.70 | 115 |

* Values were disguised to make the factorial design less transparent by adding or subtracting $5 from the spread.

the nine loss items. The presentation orders for problems within a domain were randomized for each subject and session. All subjects responded to the same 18 gambles.

Thirty-six paid junior and senior undergraduates, familiar with decision analysis from coursework, participated. They were randomly assigned to conditions (19 to PE-CE, 17 to CE-PE), and run in groups. One week after their first session, they participated in a second procedurally equivalent session, using the alternate response mode and individualized forms that reflected their first-stage responses. Each session required about 30 minutes.

### *Results*

We organize our bias results into two parts: we first describe the general pattern for each response mode order, and then the effects of risk attitude.

### *Observed Bias*

*CE-PE Condition.* Here bias is the difference between the second stage PE response, and the Initial Probability presented to respondents in the first stage CE question. The observed bias, shown in Figure 1, is significantly positive (mean = 0.073, $p < 0.001$). There is a marked effect of Initial Probability, with 0.3 items showing the most positive bias (mean = 0.175), 0.5 items show a more moderate bias (mean = 0.065), and 0.7 items a slight negative bias (mean = −0.019). There is no systematic effect of Spread. The lines appear roughly parallel, and there is little difference between the patterns for gains and losses.

These effects are confirmed by a repeated measures ANOVA, that uses, as within-subjects factors, Domain, Spread (nested within domain), Initial Probability, and their interactions. Only one experimental factor, Initial Probability ($F[2,253] = 60.83$, $p < 0.001$) is significantly related to bias. Since there is no systematic effect of Spread, nor an interaction, the lines in Figure 1 are essentially parallel. In sum, despite the differences

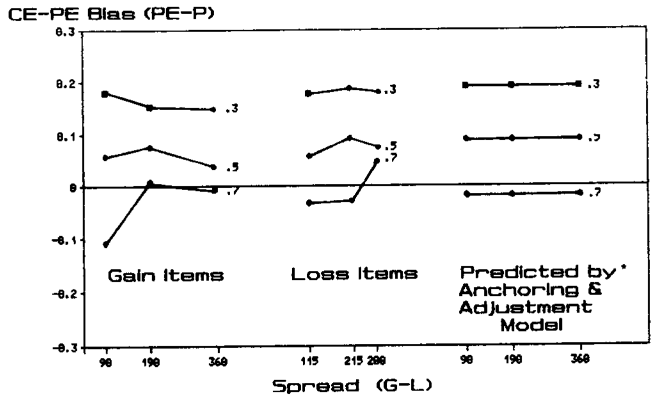



FIGURE 1. CE-PE Bias Experiment 1.
* Generated by using estimated parameters in Table 5 and assuming risk neutrality.

in these procedures and stimuli, we essentially replicate Hershey and Schoemaker's main result, that probability responses during the second stage are, on average, larger than the probability shown during the first stage. However, we also find that the smaller the Initial Probability, the more positive the bias.

*PE-CE Condition*. In this response mode order, bias is the difference between the second-stage CE response and the sure amount $S$ presented in the first-stage PE question. The observed bias shows a more complicated pattern than that in the CE-PE case, with effects of Initial Probability, Spread, and their interaction (Figure 2).

An ANOVA, identical to the preceding analysis, confirms these effects. There are significant main effects of Spread ($F[4,288] = 2.70$, $p < 0.05$), and Initial Probability ($F[2,288] = 67.01$, $p < 0.001$). Bias is large and positive for 0.3 items (mean = 32.38), smaller but positive for 0.5 items (mean = 10.65), and negative for 0.7 items (mean = −11.17). The overall mean bias is again significantly positive (mean = 10.62, $p < 0.001$). The Spread by Initial Probability interaction ($F[8,288] = 2.96$, $p < .01$) produces the "fan" effect seen in Figure 2. The larger the Spread, the greater the difference in bias for between probability levels. There is no effect of domain, or any other significant interactions.

In sum, the observed bias in this condition is a function of three factors: Initial Probability (the smaller the probability, the more positive the bias), Spread (the larger the

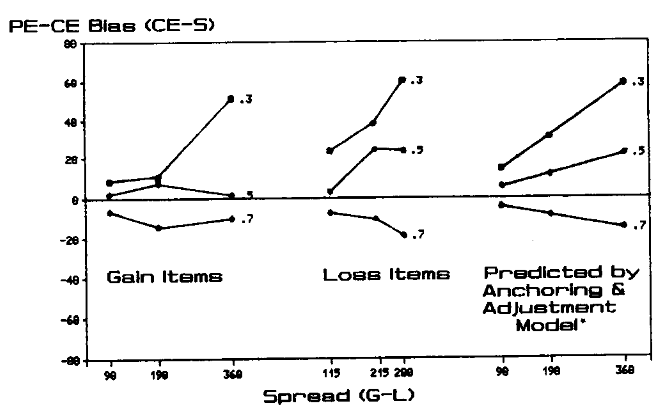



FIGURE 2. PE-CE Bias, Experiment 1.
* Generated by using estimated parameters in Table 5 and assuming risk neutrality.

spread, the larger the bias), and their interaction (the observed fan effect). This pattern, while complex, essentially replicates and extends Hershey and Schoemaker's results. They find that the size of bias, for 0.5 items, increases as the nonzero outcome increases in size. Our results extend this pattern by identifying an influence of the reference probability, and its interaction with the difference between outcomes.

### *Risk Attitude and Bias*

We now turn to the relationship between risk attitude, as indicated by first stage responses, and bias. Recall that Hershey and Schoemaker report that initial risk attitude and response mode order have a strong interactive effect on bias: for PE-CE, only risk-averse subjects show systematic bias, while for CE-PE, only risk-seeking subjects show systematic bias.

Our results provide only partial confirmation of the Hershey and Schoemaker interaction. We confirm that risk aversion leads to a positive bias in PE-CE (CE $> S$), and risk seeking to a positive bias in CE-PE (PE $> p$) (Table 2). However, contrary to Hershey and Schoemaker's description of their data, Table 2 shows that there are strong and systematic effects for all risk attitudes, but that the mean effect varies across risk attitudes. In particular, a negative bias is found for risk aversion in CE-PE, and for risk seeking in PE-CE. Thus, increased risk aversion, as measured by first stage responses, creates a more positive bias in the PE-CE response order, but a more negative bias in the CE-PE response order.

These results are confirmed by an ANOVA like those above, but with risk attitude and its interactions added as factors. Risk attitude has a significant effect in both PE-CE ($F[2,288] = 6.77$, $p < 0.001$) and CE-PE ($F[2,253] = 19.53$, $p < 0.001$). While the effect of risk attitude is highly significant, it is apparently unrelated to other effects: none of its interactions with other factors are statistically significant.

### *Discussion*: *Experiment* 1

The results of Experiment 1 replicate and extend one of Hershey and Schoemaker's key results: a positive bias in both PE-CE and CE-PE. We can compare our 0.5 items

TABLE 2
*Inconsistencies by Risk Attitude**

| | | Experiment 1 | |
|---|---|---|---|
| | | PE-CE bias (CE − $S$) | CE-PE bias (PE − $P$) |
| Risk Attitude at Time 1 | Averse | 18.66 | −0.0352 |
| | Neutral | 1.60 | 0.0789 |
| | Seeking | −12.51 | 0.1572 |

| | | Experiment 2 | |
|---|---|---|---|
| | | PE-CE bias (CE − $S$) | CE-PE bias (PE − $P$) |
| Risk Attitude at Time 1 | Averse | 13.74 | 0.0497 |
| | Neutral | −4.03 | 0.0745 |
| | Seeking | −17.30 | 0.1429 |

* All differences within a response mode are significant by a Duncan's test, $\alpha = 0.05$.

with theirs to examine possible differences in the degree of bias. A recent experiment by Byrd, de Neufville and Delquie (1987) also uses a variant of Hershey and Schoemaker's procedure and provides another comparison: for CE-PE, our 0.5 items have a mean bias of 0.065, very close to the value of 0.068 found by Hershey and Schoemaker, and the 0.078 found by Byrd et al. To facilitate comparisons across studies and items, we normalize PE-CE bias by computing it as a percentage of spread $(G - L)$. Here again the results of the three studies are comparable: our subjects average a 6.9% bias across the 0.5 items, Hershey and Schoemaker's a 7.9% bias, and Byrd et al. 7.6%. Thus, despite differences in subjects, scenarios, and items, there appears to be a high degree of comparability in these three studies.

However, these results provide a more detailed description of the bias. Bias depends on Initial Probability in both response mode orders, and in PE-CE, also on Spread, and the interaction of Spread with Initial Probability. Under certain conditions, such as for gambles with an Initial Probability of 0.7, the bias is even negative.

Thus, the picture that emerges extends the pattern of bias developed by Hershey and Schoemaker, placing further demands upon possible explanations for the bias. The risk attitude results, however, tentatively suggest that Hershey and Schoemaker's description may need modification. Risk attitude appears to affect the size of the bias, but not eliminate it. Why do our risk attitude results differ from Hershey and Schoemaker's? One possibility is that explicit anchors were not presented, allowing subjects to choose their own, a possibility we examine in Experiment 2. Whatever the reason, the form of the effect of risk attitude on bias is important, since Hershey and Schoemaker eliminate several alternative explanations, based in part upon this finding. These include mechanisms such as random error, stochastic utility functions, and anchoring and adjustment. Thus, if the pattern of bias is different than that provided by Hershey and Schoemaker, other explanations may warrant reconsideration.

## Experiment 2

### *Overview*

In Experiment 2, we wish to replicate our previous findings, particularly those concerning the relationship between risk attitude and bias, in the presence of explicit anchors. Like Hershey and Schoemaker's procedure, we first ask subjects if their response is above, below, or equal to some potential response value. We could present two kinds of potential response values, either (1) the risk neutral response (e.g., 0.3 for 0.3 gambles) or (2) anchors fixed at certain levels across gambles (e.g., 0.3 for half the items, 0.7 for the rest). We chose the latter since it allows us to examine possible effects of anchoring (Slovic and Lichtenstein 1971; Tversky and Kahneman 1974). Also, the former confounds anchor with two factors we wish to manipulate: Initial Probability in PE, and Expected Value in CE. In addition, we separately manipulate Spread and Expected Value, which were partially confounded in Experiment 1.

### *Method*

Ninety-nine junior and senior business majors participated for course credit. They were familiar with decision analysis and elicitation methods. The order of items within a response mode was randomly varied across subjects. Through randomization, 48 subjects were assigned to CE-PE, and 51 to PE-CE.

In both response modes, all subjects completed a new set of 18 items, created by the factorial combination of three levels of Initial Probability (0.3, 0.5, and 0.7), two levels of Spread ($100 and $200), and three levels of Expected Value ($150, $250 and $350) (Table 3).

TABLE 3
*Stimuli for Experiments 2 and 3*

| Gamble | $S$ | $G$ | $L$ | $p$ | Spread |
|---|---|---|---|---|---|
| 1 | 150 | 220 | 120 | 0.3 | 100 |
| 2 | 150 | 200 | 100 | 0.5 | 100 |
| 3 | 150 | 180 | 80 | 0.7 | 100 |
| 4 | 250 | 320 | 220 | 0.3 | 100 |
| 5 | 250 | 300 | 200 | 0.5 | 100 |
| 6 | 250 | 280 | 180 | 0.7 | 100 |
| 7 | 350 | 420 | 320 | 0.3 | 100 |
| 8 | 350 | 400 | 300 | 0.5 | 100 |
| 9 | 350 | 380 | 280 | 0.7 | 100 |
| 10 | 150 | 290 | 90 | 0.3 | 200 |
| 11 | 150 | 250 | 50 | 0.5 | 200 |
| 12 | 150 | 210 | 10 | 0.7 | 200 |
| 13 | 250 | 390 | 190 | 0.3 | 200 |
| 14 | 250 | 350 | 150 | 0.5 | 200 |
| 15 | 250 | 310 | 110 | 0.7 | 200 |
| 16 | 350 | 490 | 290 | 0.3 | 200 |
| 17 | 350 | 450 | 250 | 0.5 | 200 |
| 18 | 350 | 410 | 210 | 0.7 | 200 |

Instructions were identical to those used in Experiment 1 with one change: prior to responding, subjects were asked if their response would be greater, equal or less than a certain value, or anchor, approximating Hershey and Schoemaker's procedure. For each response, subjects saw either a Low anchor or High anchor. For PE, Low anchors ranged from 0.2 to 0.3, and High anchors from 0.7 to 0.8. For CE, Low anchors were equal to $L + 0.25(G - L)$, and High anchors to $L + 0.75(G - L)$. Each subject received nine High and nine Low anchors in each response mode, producing a $2 \times 2$ factorial by crossing anchor position (High, Low) in the first stage, with anchor position (High, Low) in the second. Thus, there were four possible combinations of anchors across subjects and the two response modes for a given gamble: Low–Low, Low–High, High–Low, or High–High. This design allows us to determine whether the mere presence of an anchor, or its position relative to the response, affects the resulting bias.

*Results*

We organize the results by first examining the effects of gamble characteristics, including Expected Value, on bias, then the relationship between risk attitude and bias, and finally the effects of Anchor Position.

*Gamble Characteristics and Bias*

The overall results of the first study are replicated for both response mode orders (Figures 3 and 4). These effects are confirmed by an ANOVA similar to those in Experiment 1, but with Expected Value and Anchor Position included as factors. In CE-PE, there is the same significant effect of Initial Probability ($F[2,768] = 221.59$, $p < 0.001$) and no effect of Spread. In PE-CE, there are effects of Initial Probability ($F[2,819] = 104.68$, $p < 0.001$), Spread ($F[1,819] = 8.33$, $p = 0.002$), and a marginally significant Initial Probability by Spread interaction ($F[2,819] = 1.98$, $p = 0.102$). Again, the effect of probability increases for larger levels of Spread. There is no significant effect of Expected Value in CE-PE, but a small and curious effect in PE-CE ($F[2,819] = 4.92$, $p = 0.030$), in which gambles with EV = 150 have greater bias (mean = 8.9) than those with EV = 250 (mean = 3.8) or EV = 350 (mean = 4.4).

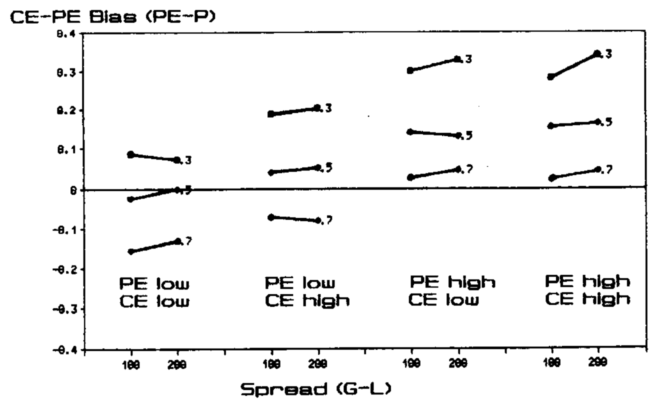



FIGURE 3. CE-PE Bias by Anchor Condition, Experiment 2.

*Risk Attitude and Bias*

The relationship between risk attitude and bias is similar to that in Experiment 1 (Table 2). In PE-CE, bias increases with the degree of risk aversion, but decreases in CE-PE. First stage risk attitude is significantly related to bias in ANOVAs for both CE-PE ($F[2,768] = 6.96$, $p < 0.001$) and PE-CE ($F[2,819] = 6.89$, $p < 0.001$). In addition, risk attitude again has no significant interaction with other experimental factors in CE-PE. However, in PE-CE there is a significant interaction with Spread ($F[1,819] = 4.03$, $p = 0.038$). The only notable difference in these results and those of Experiment 1 is the slight positive bias for risk aversion in CE-PE. However, the trend remains the same: increased risk aversion, as measured by first stage responses, creates a more positive bias in the PE-CE response order, but a more negative bias in the CE-PE response order.

*Anchor Position and Bias*

Anchor Position has a large and significant effect on bias. To test the effects of anchors, we created a factor with four levels, representing the possible anchor combinations across response modes: Low–Low, Low–High, High–Low, or High–High. Anchors have highly significant effects in both CE-PE ($F[3,768] = 92.40$, $p < 0.001$) and PE-CE ($F[3,819] = 20.18$, $p < 0.001$). As seen in Figures 3 and 4, the mean bias ranges from significantly negative in the Low–Low condition, to highly positive in the High–High condition, with

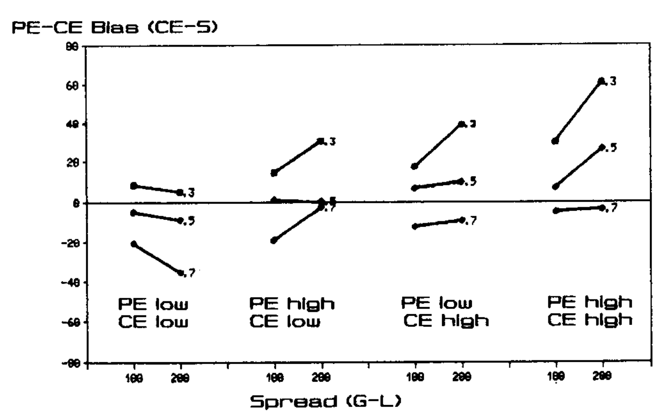



FIGURE 4. PE-CE Bias by Anchor Condition, Experiment 2.

the High–Low and Low–High conditions in between. In CE-PE, there are no significant interactions of other factors with anchor, but in PE-CE there is a significant interaction with Spread ($F[3,819] = 4.93$, $p = 0.032$).

*Discussion: Experiment 2*

Experiment 2 largely replicates the results of the first study, and confirms part of the picture provided by Hershey and Schoemaker. First we replicate their large and systematic bias and extend it to more complex stimuli with two nonzero outcomes. We have also seen that factors such as initial probability and spread affect bias, depending upon the order of the response modes. Thus the existence of the bias is reaffirmed across new stimuli, subject populations, and scenarios. In addition, we have seen that anchors can have a substantial effect on the pattern of bias.

The major discrepancy between our results and Hershey and Schoemaker's concerns the relationship between risk attitude, response mode order, and bias. Our results suggest that this relationship is not necessarily an interaction (i.e., existing for some combinations of risk attitude and response mode order, but not others). Instead, the mean bias appears to be a function of the degree of first-stage risk-attitude: in PE-CE, the more risk averse the first-stage response, the more positive the bias; in CE-PE the more risk averse the first-stage response, the more negative the bias. Because the supposed interactive nature of this relationship is used by Hershey and Schoemaker to eliminate some potential explanations for bias, its status is important.

Several explanations for this discrepancy are possible: we used a wider range of stimuli, for example. However, other leading candidates might be our use of a larger sample of items and parametric analyses that increase the statistical power of our tests. The critical tests for asymmetry in Hershey and Schoemaker's research are based on relatively small samples (an average of 13 for the trinomial tests of asymmetry, p. 1220), because few subjects displayed certain risk attitudes in some response modes. In contrast, the means in Table 2 reflect an average of 110 observations for Experiment 1, and 297 observations for Experiment 2. Further, in a more recent study Schoemaker and Hershey (1988) also found bias for all risk attitudes. Whatever the cause, however, the picture that emerges here suggests a much simpler explanation of the relationship between risk attitude and bias.

*Risk Attitudes, Bias and Error*

Recall that risk attitude is defined using a subject's first-stage response. Because it is based on a single response, this measurement of risk attitude could reflect two factors: a "true" risk preference, and random error (i.e., $\mathrm{PE}_{\mathrm{observed}} = \mathrm{PE}_t + \epsilon_p$, $\mathrm{CE}_{\mathrm{observed}} = \mathrm{CE}_t + \epsilon_c$, where $E[\epsilon_p] = E[\epsilon_c] = 0$). Hershey and Schoemaker (p. 1222) show that second-stage responses are a function such errors. We adapt their equations here for an EU maximizer in the PE-CE case. The probability presented in a stage-two CE gamble is the $\mathrm{PE}_{\mathrm{observed}}$ from stage one, and the second-stage CE response would be

$$\mathrm{CE}_2 = U^{-1}[(\mathrm{PE}_t + \epsilon_p)U(G) + (1 - \mathrm{PE}_t - \epsilon_p)U(L)] + \epsilon_c. \tag{1}$$

Because $\mathrm{PE}_t = p$, this can be rewritten as

$$\mathrm{CE}_2 = U^{-1}[\mathrm{EU} + \epsilon_p(U(G) - U(L))] + \epsilon_c. \tag{2}$$

Since $U(G) - U(L) > 0$, the first-stage error is positively related to the second-stage response, and therefore to bias. Consider a risk neutral decision maker (i.e., EU = EV). First-stage PE responses with positive errors would be classified as risk averse, and those with negative errors as risk seeking, producing the symmetric effect seen in Table 2. First-stage risk aversion (positive error) leads to positive bias, and first-stage risk seeking (negative error) to negative bias. A similar analysis for CE-PE also yields a positive relationship between first-stage error and bias. Again a symmetric effect is produced: first-stage risk seeking (positive error) leads to positive bias, and first-stage risk aversion (negative error) to negative bias. Note that this pattern is exactly that shown in our data, suggesting that error may well account for the observed risk attitude effect. However, it should also be noted that this misclassification does not explain the other systematic effects found in the bias (e.g., the influence of Initial Probability).

Of course the observed responses in these experiments are a function of *both* true risk preferences (which are not necessarily risk neutral) and random error. The ability of classification errors to produce the observed effect of risk attitude depends on the relative impact of these two factors. Neither Hershey and Schoemaker, nor we, have direct assessments of the reliability of these measures. Yet examination of the first-stage risk premiums and their standard deviations suggests that there is a substantial potential for misclassification due to error. For example, for a given gamble in Experiment 1, the mean first-stage CE response was only 0.42 standard deviations from the gamble's EV. Thus, it is quite possible that random error alone could produce a substantial "risk attitude" effect.

*Explanations*

What do our results imply for explanations of the discrepancy between CE and PE? It appears that they remove one constraint: the interactive relationship between risk attitude, response mode order, and bias. Our results show a more symmetric effect of first-stage "risk attitude" that can be largely accounted for by random error. However, they impose another set of constraints: the systematic effects of Initial Probability and Spread and their interaction with response mode order. We will assess the status of possible explanations more carefully in a later section.

## Experiment 3: Process Analysis

*Overview*

One consequence of the first two experiments is that other explanations for the bias may need to be reconsidered. Further, it is possible that several different mechanisms may underlie these results. Goldstein and Einhorn (1987) have suggested that several mechanisms might underlie another common response mode anomaly: preference reversals. Similarly, Hershey and Schoemaker suggest that the observed pattern of bias might arise from multiple sources.

Examining this possibility more carefully requires additional forms of data. In this study we collect process measures that might allow us to identify some of the processes underlying bias by looking at intermediate behaviors such as patterns of information search, response latencies, and verbal reports. For a more complete review of process techniques see Payne, Braunstein and Carroll (1978), and Russo (1978).

*Method*

Because our emphasis in this study is on the underlying cognitive processes, we focus our attention on the response mode order with the simplest pattern of bias, CE-PE. Specifically, we examine concurrent verbal reports, and collect data on information acquisitions using a novel technique. Gambles are presented to subjects on a computer display, illustrated in Figure 5. The values of the gamble elements are not visible initially, but are "hidden" behind labeled boxes. Values are revealed by moving a cursor into each cell using a hand-controlled pointing device called a *mouse*. Subjects give responses by moving a pointer along the continuous scale using the mouse. The software records the

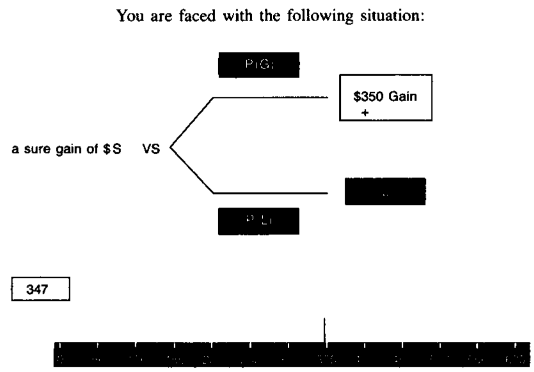



FIGURE 5. Computer Display of Assessment Gamble.

order and length of time for which each piece of information is accessed, as well as any intermediate responses made on the scale. Further details about the *Mouselab* system are found in Johnson, Payne, Schkade and Bettman (1988).

Twenty-four paid juniors, seniors, and masters students participated in two procedurally equivalent individual sessions, held at least three days apart. They were familiar with utility theory and decision analysis from course work. After a brief presentation familiarizing them with the mouse, subjects received the same instructions and items used in Experiment 2, but without anchors. They then received training in generating think-aloud protocols (Ericsson and Simon 1984), and answered the last four items while generating verbal reports.

We coded the transcribed verbal protocols for: (1) explicit use of arithmetic operations such as multiplying probabilities and payoffs (e.g., "0.5 times $360 is $180 . . ."), (2) use of an adjustment strategy as indicated by a series of incremental responses (e.g., "say $360, no, let's say $370 . . ."). In addition, we also noted any explicit mention of an anchor (e.g., "let's start with $360 . . ."), although these were infrequent. One-third of the protocols were coded by a rater naive to our hypotheses, who showed substantial agreement (93% of sessions assigned to the same strategy).

*Results*

*Process Analysis*. The protocols showed strong individual differences in strategy. Subjects tended to consistently employ either an expected value or an heuristic strategy that contains elements of both anchoring and adjustment and PE reframing. We found no case in which a subject clearly calculated expected value on one trial and then switched to heuristic strategy on another. Based upon these data, we categorized people into two groups:

1. An Expectation group (9 subjects) who explicitly multiplied or divided, calculating expected value consistently for either the Probability or Certainty equivalence task, and
2. A Heuristic group (15 subjects) who did not explicitly calculate, but instead mentioned intermediate answers that were incremented toward the final response.

While not all trials could be classified, there is enough evidence across the four trials to clearly assign each session to a strategy group. The strategy of the Heuristic group is illustrated by the following CE protocol:

> Okay a 70% chance at a $410 gain
> a 30% chance at a $210 dollar gain
> About a 200 dollar difference

so I would probably take . . .
not around 310
I would probably take much less . . .
let's see 310 would be halfway in between
so we would take something like 275
something close to that because the
probability of gain is so high
274.

Subjects in this group also showed some tendency to reframe PE alternatives, as described by Hershey and Schoemaker. While the tendency to reframe is less frequent than anchoring and adjustment, it appears in several PE protocols:

390, 190, 240 . . .
390 versus 240 . . .
that's a pretty big amount of difference . . .
240 and 190 . . . $50 difference . . .
so I think I'd probably take the extra $50, so . . .
*p* would have to be pretty high . . .
I'd say 78 to risk the 190 . . .

Process measures from the computer logs support our categorization of the protocols.[2] Consistent with the notion of adjustment, the Heuristic group spent more time adjusting the response scale, and made more adjustments, than the Expectation group. In addition, there is an interesting difference in the way the two groups search for information in the PE task. The Expectation group made frequent transitions between the boxes containing $G$ and $L$, and $S$ and $L$, but rarely moved between $S$ and $G$ (only 16% of their transitions fall into this latter category). This is consistent with the observation that these subjects calculate expected value using the ratio $(S - L)/(G - L)$, which would not produce transitions between $S$ and $G$. In contrast, in the Heuristic group, 38% of transitions fall into this category, which is significantly higher than the Expectation group ($p < 0.05$ when tested by an ANOVA similar to those in Experiments 1 and 2). Further, the correlation, across subjects, between the frequency of this transition and the amount of bias, is significantly positive ($r[1,394] = 0.309$). Analogous effects in the CE response mode are consistent but fail to reach statistical significance.[3] Thus, multiple measures of underlying cognitive processes support the notion that the two groups use different strategies: one employing explicit multiplication, and another employing heuristics such as anchoring and adjustment and reframing.

*Outcome Analysis*. An important question suggested by these differences is if strategy, as revealed by the protocols, is related to bias. Figure 6 shows the observed bias for each strategy group, as a function of Initial Probability and Spread. The Expectation group shows almost no bias, and the overall pattern of bias (which replicates that in Experiments 1 and 2) is generated almost exclusively by the Heuristic group.

These observations are confirmed by a repeated measures ANOVA similar to those in Experiments 1 and 2, but including strategy as a between subjects factor. Not only is there a significant main effect of Strategy ($F[1,21] = 6.04$, $p < 0.01$), but a sizable

[2] Given concerns over the possible effects of verbalization, we conducted an analysis comparing protocol to nonprotocol trials. In addition to bias, we examined several process measures such as time required to answer each item, and various patterns of information acquisition. None of these measures showed significant differences. Details of this analysis are reported in Russo, Johnson, and Stephens (1988).

[3] Because of computer failure, about $\frac{1}{3}$ of the observations from this group were lost for several of the process measures, perhaps contributing to this null result. Conclusions about lack of differences, therefore, are tentative.

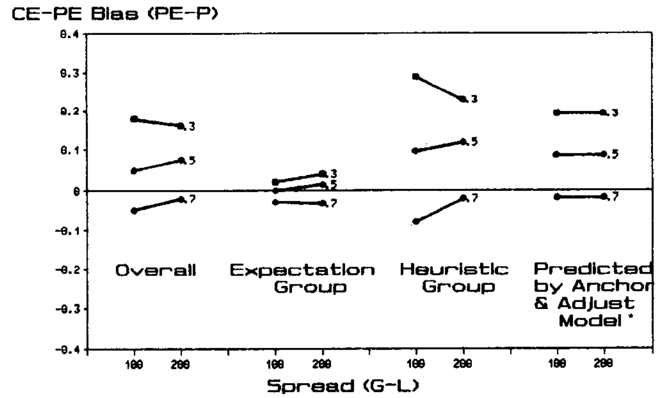



FIGURE 6. Experiment 3: Observed Bias by Strategy.

* Generated by using estimated parameters from Experiment 1 in Table 5 and assuming risk neutrality.

interaction with Initial Probability ($F[2,394] = 20.97, p < 0.001$). An analysis of simple effects shows that Initial Probability is strongly related to bias for the Heuristic group ($F[2,387] = 49.10, p < 0.001$), but not for the Expectation group ($F[2,287] = 1.43, p > 0.50$).

*Discussion*: *Experiment 3*

The protocols, along with other process tracing measures, allow us to identify two different response strategies, each of which has a different pattern of bias. Only those subjects who are identified, through verbal reports, as using heuristics, demonstrate significant bias. Thus, the protocols suggest that bias results from the use of certain heuristic strategies and gives tentative evidence that these may include aspects of both anchoring and adjustment and reframing.

## Models for Bias

In this section we review the status of two possible explanations for inconsistencies between PE and CE in light of the results of Experiments 1–3. These results are summarized in Table 4, which shows a consistent pattern across experiments. These explanations, PE reframing and anchoring and adjustment, were previously examined by Hershey and Schoemaker (1985).

*PE Reframing*

The reframing model suggests that bias results primarily from a shift toward risk aversion in PE (due to reframing), assuming that "real" utility functions are concave for gains and convex for losses (Hershey and Schoemaker 1985, pp. 1224–1226). Hershey and Schoemaker go on to show that this reframing leads to positive bias. This explanation also predicts both the ordering of mean biases across Initial Probability levels (bias most positive with 0.3 gambles, least positive with 0.7 gambles), and in PE-CE, the interaction of Initial Probability with Spread in our results (Table 4). However, it does not account for the negative bias we observed for 0.7 gambles in all three experiments. This model also implies that bias will be an increasing function of expected value. However, the only significant effect of expected value was a small effect in the opposite direction in the PE-CE condition of Experiment 2 (Table 4). This result should be interpreted cautiously, since we manipulated expected value over a relatively small range, although a similarly small manipulation of spread did have significant effects on PE-CE bias. Finally, as

TABLE 4

*Summary of Results*[1]

| | $p$ | Spread | $p$ by Spread | EV | Risk Attitude Effect[2] | Risk Attitude Interaction[3] |
|---|---|---|---|---|---|---|
| CE-PE | | | | | | |
| Experiment 1 | −*** | ns | ns | N/A | −*** | no |
| Experiment 2 | −*** | ns | ns | ns | −*** | no |
| Experiment 3 (Heu) | −*** | ns | ns | ns | −** | no |
| Experiment 3 (Exp)[4] | ns | ns | ns | ns | ns | no |
| PE-CE | | | | | | |
| Experiment 1 | −*** | +** | +*** | N/A | +*** | no |
| Experiment 2 | −*** | +** | +* | ?**[5] | +*** | no |

*$p < 0.10$.
**$p < 0.05$.
***$p < 0.01$.

[1] Entries show direction of effect on bias. For example, "−" means bias decreases as probability increases.
[2] "+" means bias increases with the degree of risk aversion.
[3] "No" means that Hershey and Schoemaker's interactive effect of risk attitude and response mode order was not found.
[4] No bias predicted.
[5] A nonmonotonic relationship was found.

described by Hershey and Schoemaker, the model predicts that response mode order and risk attitude have an interactive effect on bias. In contrast, we observe a more symmetric effect of risk attitude, that may be due largely to random error. Thus, while PE reframing probably occurs, this explanation cannot by itself completely account for the bias observed in our experiments.

### *Anchoring and Adjustment*

Hershey and Schoemaker use a mathematical model to operationalize anchoring and adjustment, in which responses are a weighted average of the expected value response (anchor) and the "true" utility maximizing response. The evidence provided by these three studies suggests that a generalized version of the model may deserve reconsideration: Experiment 2 showed that anchors are sufficient, if not necessary, to cause bias, while Experiment 3 provided protocol evidence of an adjustment process. We generalize their model in two ways. First, since our gambles involved two nonzero outcomes, we reformulate the models to include both $G$ and $L$. Second, the protocols suggest that decision makers seem to generate their own ball-park guess or starting point. This does not always correspond to the expected value. Equations (3) and (4) show Hershey and Schoemaker's (pp. 1226–1227) formulations for gambles with one nonzero outcome and expected value anchors:

For CE-PE:

$$\mathrm{PE}_2 = w_p\left\{\frac{U[w_c U^{-1}(pU(G)) + (1 - w_c)pG]}{U(G)}\right\} + (1 - w_p)\left\{\frac{w_c U^{-1}(pU(G)) + (1 - w_c)pG}{G}\right\} \quad (3)$$

and for PE-CE:

$$\mathrm{CE}_2 = w_c\left\{U^{-1}\left[U(G)\left[w_p\left(\frac{U(pG)}{U(G)}\right) + (1 - w_p)p\right]\right]\right\} + (1 - w_c)G\left\{w_p\left(\frac{U(pG)}{U(G)}\right) + (1 - w_p)p\right\} \quad (4)$$

where $w_p$ and $w_c$ are the relative weights given to the utility maximizing response, and $U$ and $U^{-1}$ refer to a utility function and its inverse. To incorporate both $G$ and $L$, and self-anchoring, let $A_p$ and $A_c$ represent anchors. For example, a single-stage PE response now becomes

$$\mathrm{PE}_1 = w_p\left[\frac{U(S) - U(L)}{U(G) - U(L)}\right] + (1 - w_p)A_p \quad (5)$$

and a single-stage CE response becomes

$$\mathrm{CE}_1 = w_c U^{-1}[pU(G) + (1 - p)U(L)] + (1 - w_c)A_c. \quad (6)$$

Reformulating equations (3) and (4) to include these modifications yields equations (7) and (8):

For CE-PE:

$$\mathrm{PE}_2 = w_p\left\{\frac{U[w_c U^{-1}[pU(G) + (1 - p)U(L)] + (1 - w_c)A_c] - U(L)}{U(G) - U(L)}\right\} + (1 - w_p)A_p \quad (7)$$

and for PE-CE:

$$\mathrm{CE}_2 = w_c U^{-1}\left\{[U(G) - U(L)]\left[w_p\left(\frac{U(S) - U(L)}{U(G) - U(L)}\right) + (1 - w_p)A_p\right] + U(L)\right\} + (1 - w_c)A_c. \quad (8)$$

Responses are now expressed as a weighted average of the utility maximizing response and the self-generated anchor. If adjustment is insufficient, $w_p$ and $w_c$ will be less than 1, and anchors will bias responses away from utility maximization. The pattern of bias produced by this model depends critically on the anchors used by subjects, the degree of insufficient adjustment, and to a lesser extent, the "true" risk attitude.[4] Consistent with the results of Experiment 2, the model predicts a more positive bias for higher anchors, a more negative bias for lower anchors, and except for an interaction with spread in PE-CE, that the effects of anchors are independent of other factors. Moreover, the model predicts the qualitative pattern of bias observed across experiments, summarized in Table 4: an effect of probability in CE-PE (see Figure 1), and of probability, spread, and their interaction in PE-CE (see Figure 2).

More precise predictions depend upon the specific parameter values. Using equations (5) and (6) and only the first-stage PE and CE responses, the weights and anchors can be estimated. First, assume that $A_p$ is constant across gambles (i.e., subjects tend to anchor at the same point on the scale). Similarly, define $A_c$ as the proportion of the distance between $G$ and $L$ at which subjects anchor (i.e., in equation (5) replace $A_c$ with the expression $[A_c(G - L) + L]$). Thus, both anchors range from 0 to 1. Finally, leave the weights unconstrained, and assume risk neutral preferences. Additional details can be found in the Appendix.

[4] Simulations provided predictions for both the self-generated anchor and PE reframing models.

Nonlinear least squares estimations yield the values shown in Table 5. Using the estimates derived from first-stage responses in Experiment 1, the predictions of equations (7) and (8) for the gambles in Experiment 1 are shown in Figures 1 and 2 (predictions are the same for gains and losses), and for Experiment 3 in Figure 6. The pattern of bias in both response mode orders is fit well, including the net positive bias. To test model fit, we regressed the predictions against the observed mean bias for the eighteen gambles, yielding the $R^2$ values shown in Table 5. Similar results for Experiments 2 and 3 are also shown in Table 5.

In all three experiments, the estimates of $w_c$ and $w_p$ are less than 1, implying insufficient adjustment. For CE, estimated anchors are essentially in the middle of the scale across studies, and PE anchors are about halfway between the middle and top of the scales. While it is plausible that subjects might anchor on a middle value, as in CE, the psychological motivation for choosing an anchor at three-quarters of the scale range is less apparent. One explanation is that PE subjects sometimes used the top of the scale, and sometimes the middle, producing an intermediate value. The more extreme anchors in PE could be related to results in Schoemaker (1987), who suggests that anchoring effects may be stronger in PE than CE.

The estimations also reflect the experimental manipulation of anchors in Experiment 2 and the process analysis of Experiment 3. In Experiment 2 the high anchor condition shows higher estimated anchors than the low anchor condition. In Experiment 3 only the Heuristic group shows insufficient adjustment. This congruence between the estimates generated by the model, and both the experimental manipulation and process tracing results supports the model's characterization of the data.

## General Discussion

The two explanations we have considered affect different stages of judgment: PE-CE reframing concerns the encoding of the outcomes, and anchoring and adjustment describes the evaluation of the options. Further, Goldstein and Einhorn (1987) suggest that the bias may result from expressing an evaluation using different response scales. These mechanisms could act independently, and therefore all play a role in explaining bias.

Furthermore, the models are quite similar in some respects. For example, the reframing model produces predictions that have many of the same ordinal characteristics as those for the anchoring and adjustment model (e.g., the ordering of bias across reference probabilities). Given these similarities, it may be difficult to isolate the locus of bias as occurring in one stage or another. The mapping between these models needs to be more formally developed, but the surprising similarity of predictions emphasizes the potential of carefully

TABLE 5

*Estimated Anchors and Adjustment Parameters: The Self-Generated Anchor Model**

| | $W_c$ | $W_p$ | $A_c$ | $A_p$ | $R_c^2$ | $R_p^2$ |
|---|---|---|---|---|---|---|
| Experiment 1 | 0.75 | 0.63 | 0.53 | 0.73 | 0.86 | 0.86 |
| Experiment 2 | | | | | | |
| Stage 1 anchor High | 0.59 | 0.49 | 0.68 | 0.84 | 0.90 | 0.88 |
| Stage 1 anchor Low | 0.71 | 0.61 | 0.37 | 0.65 | 0.88 | 0.68 |
| Experiment 3 | | | | | | |
| Heuristic Group | 0.74 | — | 0.54 | — | — | 0.91 |
| Expectation Group | 1.02 | — | —** | — | — | 0.69 |

* All estimations are based on Stage 1 responses to minimize confounding with bias (see appendix).

** In the Expectation Group, the anchor was given essentially no weight, and thus estimates of the anchor become unstable, and irrelevant.

designed experiments and process data in isolating the relative contributions of possible sources of bias. In particular, methods that could separate reframing from anchoring would seem useful.

To summarize, this research makes three contributions: First, by extending the range of stimuli we have provided a broader characterization of the possible biases observed in utility assessment. Second, we have demonstrated through an experimental manipulation that anchoring *can* cause bias and provided process tracing evidence that the use of heuristic strategies such as anchoring and reframing may be involved in generating bias. Third, we generalize Hershey and Schoemaker's weighted-average model to include self-generated anchors. Such a model seems a viable explanation, since it captures several aspects of the data from the three experiments. However, the final explanation for bias will probably involve multiple influences, including perhaps effects such as response compatibility (Tversky, Slovic and Sattath 1988) that we have not discussed here.

While we have just begun the process of depicting the strategies adopted by decision makers in utility assessment, this research also raises several more general issues. First, our results suggest that decomposition may not always aid judgments. It appears that subjects adopt heuristics even in these simple judgments, and that these heuristics can produce systematic biases. It therefore becomes less apparent that combining these errorful judgments will produce an overall evaluation that is superior to a holistic judgment. The value of decomposition, and the role of simplified strategies, clearly depend upon research establishing conditions under which decomposition provides suitable inputs to decision aids (for further discussion of this issue see Kleinmuntz 1988). A major shortcoming of research on this and other utility assessment biases, however, is the lack of evidence about whether they also occur in settings that are closer to decision analysis practice. Second, the application of a behavioral analysis in a utility assessment task can potentially provide insights that would improve the performance of this prescriptive procedure. The synergy between prescriptive and descriptive research benefits both. Finally, it is not apparent which response mode better reflects a decision maker's true preferences. Yet the issue of which set of preferences would make the decision maker happier is crucial to the practice of decision aiding. If we accept a constructive view of utility assessment, we are faced with attempting to decide which of several revealed preference functions is best (See Fischhoff et al. 1980 for examination of this problem.) Identifying sources of bias would be an important step toward ultimately resolving this dilemma. In the meantime, analysts are well advised to use multiple method to guard against the sensitivity of assessments to response mode biases.[5]

[5] This research was supported by a grant from the Decision, Risk, and Management Science program of the National Science Foundation to the first author. We would particularly like to thank our co-investigators on that grant, Gregory W. Fischer and Mark S. Kamlet, for many helpful discussions, and John S. Carroll, Robyn Dawes, John Hershey, John W. Payne, J. Edward Russo, Paul Schoemaker, and the anonymous reviewers, for comments on earlier drafts. We would also like to acknowledge the Engineering Psychology Program of the Office of Naval Research to Duke University for supporting the development of the software employed in Experiment 3.

## Appendix

Since anchoring and adjustment applies to just one response at a time, we can use the observed first-stage responses to estimate the parameters in equations (5) and (6). This procedure minimizes the connection between the estimates and the observed bias, since second-stage responses are not used in the estimation, and the estimates for one response mode come from an independent group of subjects (e.g., the estimates of $w_c$ and $A_c$ necessary to make bias predictions for the PE-CE group come from the CE-PE group's first-stage responses). In a single-stage response we have two variables to use in estimating the anchor and the weight: the expected utility response, and the observed response. Note that the parameters to be estimated, $w$ and $A$, have a nonlinear relationship with the responses, thus a nonlinear estimation procedure available in SAS was used.

Our first estimation assumes risk neutrality, or $U(x) = x^{\theta}$, where $\theta = 1.0$. Because the estimates are potentially sensitive to single aberrant points, a Windsorizing procedure was used to control the effects of outliers ($\pm 3\sigma$). The predictions depicted in Figures 1 and 4 assume risk neutrality.

Sensitivity analyses were also performed by estimating the parameters for a factorial design of risk averse, risk neutral, and risk seeking utility functions for losses and gains. Specifically, we varied risk attitude by setting $\theta$ to 0.5, 0.67, 0.75, 1.0 and 1.3, and allowing $\theta$ to differ for losses and gains. These transformations do not exert much influence on the estimates, and the results reported earlier (computed for a $\theta = 1.0$) are representative of the functions tested. For example, the estimated adjustment parameters across $\theta$ values for PE in Experiment 1 range only from 0.59 to 0.65.

## References

BETTMAN, J. R., *An Information Processing Theory of Consumer Choice*, Addison-Wesley, Reading, MA, 1979.

BYRD, J. G., R. DE NEUFVILLE AND P. DELQUIE, "The Difference Between Probability and Certainty Equivalent Methods of Utility Assessment," Working paper, Technology and Policy Program, MIT, 1987.

ERICSSON, K. A. AND H. A. SIMON, *Protocol Analysis*. The MIT Press, Cambridge, MA, 1984.

FARQUHAR, P. H., "Utility Assessment Methods," *Management Sci.*, 30 (1984), 1283–1300.

FISCHHOFF, B., P. SLOVIC AND S. LICHTENSTEIN, "Knowing What You Want: Measuring Labile Values," In T. Wallstein (Ed.), *Cognitive Processes in Choice and Decision Behavior*. Erlbaum, 1980.

GOLDSTEIN, W. M. AND H. J. EINHORN, "Expression Theory and the Preference Reversal Phenomena." *Psychological Rev.*, 94 (1987), 236–254.

HERSHEY, J. C., H. C. KUNREUTHER AND P. J. H. SCHOEMAKER, "Sources of Bias in Assessment Procedures for Utility Functions," *Management Sci.*, 28 (1982), 936–954.

——— AND P. J. H. SCHOEMAKER, "Probability vs. Certainty Equivalence Methods in Utility Measurement: Are They Equivalent?" *Management Sci.*, 31 (1985), 1213–1231.

JOHNSON, E. J., J. W. PAYNE AND J. R. BETTMAN, "Information Displays and Preference Reversals," *Organizational Behavior and Human Decision Processes*, 42(1988), 1–21.

———, ———, D. A. SCHKADE AND J. R. BETTMAN, "Monitoring Information Acquisitions and Decisions: Mouse Decision Laboratory Software. MouseLab User's Manual," Fuqua School of Business, Duke University, revised 1986, 1988.

KARMARKAR, U. S., "Subjectively Weighted Utility: A Descriptive Extension of the Expected Utility Model." *Organizational Behavior and Human Performance*, 21 (1978), 61–72.

KLEINMUNTZ, D. N., "Decomposition and the Control of Errors in Decision Analytic Models," Working paper #2023-88, Sloan School of Management, M.I.T., 1988.

LICHTENSTEIN, S. AND P. SLOVIC, "Reversals of Preference between Bids and Choices in Gambling Decisions." *J. Experimental Psychology*, 89 (1971), 46–55.

——— AND ———, "Response-Induced Reversals of Preference in Gambling: An Extended Replication in Las Vegas." *J. Experimental Psychology*, 101 (1973), 16–20.

MCCORD, M. AND R. DE NEUFVILLE, "Utility Dependence on Probability: An Empirical Demonstration," *J. Large Scale Systems*, 6 (1984), 91–103.

PAYNE, J. W., "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," *Organizational Behavior and Human Performance*, 22 (1976), 17–44.

———, M. L. BRAUNSTEIN AND J. S. CARROLL, "Exploring Pre-Decisional Behavior: An Alternative Approach to Decision Research," *Organizational Behavior and Human Performance*, 22 (1978), 17–44.

RUSSO, J. E., "Eye Movements Can Save the World: A Critical Evaluation and Comparison between Eye Fixations and Other Information Processing Methodologies," In J. K. Hunt (Ed.), *Advances in Consumer Research*, Association for Consumer Research, Ann Arbor, MI, 1978.

———, E. J. JOHNSON AND D. L. STEPHENS, "The Validity of Verbal Protocols," Working paper, Cornell University, 1988.

SCHOEMAKER, P. J. H., "Are Basic Risk Preferences Related Across Payoff Domains and Response Modes?" Unpublished manuscript, Center for Decision Research, University of Chicago, 1987.

——— AND J. C. HERSHEY, "A Domain-Shift Theory of Response Mode Biases in Utility Measurement," Unpublished manuscript, Center for Decision Research, University of Chicago, 1988.

SLOVIC, P. AND S. LICHTENSTEIN, "Preference Reversals: A Broader Perspective," *Amer. Economic Rev.*, 73 (1983), 596–605.

TVERSKY, A. AND D. KAHNEMAN, "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185 (1974), 1124–1131.

———, S. SATTATH AND P. SLOVIC, "Contingent Weighting in Judgment and Choice," *Psychological Rev.*, 95(1988), 371–384.