N89-19849

A METHOD FOR MODELING BIAS IN A PERSON'S ESTIMATES OF LIKELIHOODS OF EVENTS

Thomas E. Nygren and Osvaldo Morera

Department of Psychology
The Ohio State University
404C W. 17th Avenue
Columbus, OH 43210

It is of practical importance in decision situations involving risk to train individuals to transform uncertainties into subjective probability estimates that are both accurate and unbiased. We have found that in decision situations involving risk, people often introduce subjective bias in their estimation of the likelihoods of events depending on whether the possible outcomes are perceived as being "good" or "bad". Until now, however, the successful measurement of individual differences in the magnitude of such biases has not been attempted. In this paper we illustrate a modification of a procedure originally outlined by Davidson, Suppes, and Siegel [3] to allow for a quantitatively-based methodology for simultaneously estimating an individual's subjective utility and subjective probability functions. The procedure is now an interactive computer-based algorithm, DSS, that allows for the measurement of biases in probability estimation by obtaining independent measures of two subjective probability functions (S^+ and S^-) for "winning" (i.e., good outcomes) and for "losing" (i.e., bad outcomes) respectively for each individual, and for different experimental conditions within individuals. The algorithm and some recent empirical data are described.

It is argued that, if in decision situations involving substantial risk or potential loss, our goal is to train individuals to become expert decision makers, it is important to understand how people subjectively evaluate and represent uncertainties or probabilities. Decision theorists have argued for some time that any decision analysis under risk must involve the assessment of uncertainties, and that uncertainties can best be measured by subjective probabilities that represent the decision maker's degree of belief about the relevant uncertain events. The decision maker must somehow transform uncertainties into subjective probability estimates that are both accurate and unbiased. There is, however, convincing evidence that in many complex decision situations involving risk or uncertainty, people use heuristics that often introduce bias in the subjective estimation of the likelihoods of events that are relevant to the outcomes of the decision [8, 20]. They may judge the probability of an event by its representativeness of a class of events, by its availability in memory as a relevant example, or on the basis of an adjustment from a numerical anchor point.

Recently published edited volumes by Arkes and Hammond [1] and Kahneman, Slovic, and Tversky [10] reflect this new direction of study in the field of judgment and decision making research. Studies have shown that people generally do not make good probability estimates [6, 10, 11, 16, 17, 18]. They overestimate low and underestimate high probabilities and they ignore base-rate information [2, 12]; they revise opinions too conservatively [4]; they indicate excessive confidence in their judgment [7]; and they are influenced by their affective mood state [9, 15].

However, only recently have systematic research efforts investigating the cognitive mechanisms by which biases are generated been reported. For example, Nygren and Isen [15] have shown that a positive mood state can lead decision makers to exhibit "cautious optimism" in risky choice situations. They become optimistic in the sense that they tend to overestimate the likelihood of "good" events and underestimate the likelihood of "bad" events; but at the same time they exhibit a cautious shift toward risk-aversion in their actual choices.

Such findings imply the need for models that interrelate cognitive processes and judgmental biases. Wickens

[21] has argued that without an understanding of these biases in such a framework, it is difficult to predict how specific decisions are being made by individuals. But, how can such biases be quantitatively measured? Most models of decision making under risk assume that there are four basic questions that remain the focal issues in decision analysis. They are: (1) what are the possible courses of action? (2) what are the outcomes associated with these courses of action? (3) what is the utility associated with each outcome?, and (4) what is the probability associated with each outcome? Much quantitative and empirical research continues to focus on Questions 1, 2, and 3, and, in particular, the measurement of *utility* [5, 10]. This paper describes a method to take a closer look at Question 4, the measurement of the *subjective probability function*.

MATHEMATICAL MODELS

There are two leading models of risky decision making upon which this research is based, subjective expected utility (SEU) theory and Kahneman and Tversky's prospect theory [13]. In SEU theory the overall utility of a course of action or "gamble" is found by taking, for each possible outcome in the gamble, the product of the utility of the outcome multiplied by the subjective probability associated with that outcome's occurrence, and summing these terms across all outcomes. The decision maker is assumed to choose the gamble/option with the highest overall expected utility. Prospect theory proposes that the decision process is, in fact, completed in two phases, with the potential courses of action first being "framed" for the choice process. This framing often may constitute a preliminary look at the outcomes, and this look sometimes results in a simplified representation of the choice alternatives, particularly if the alternatives are complex. Following this initial phase, the alternatives are actually evaluated in a manner similar to that suggested by SEU theory, where the alternative with the highest value (utility) is chosen.

Both models take the same general form, then, in that overall preference for a course of action or gamble (G) is assumed to be a function of (a) the values or utilities of the possible outcomes and (b) the subjective probabilities (in SEU theory) or decision weights (in prospect theory) associated with these outcomes. Expressed mathematically, for a simple gamble of the form $G = (x, p; y, 1-p)$ where one obtains outcome x with probability p or outcome y with probability $1-p$, the subjective value of the gamble (G) is assumed in these models to be

$$V(G) = V(x) * S(p) + V(y) * S(1-p) \quad (1)$$

where $V(x)$ is the utility of outcome x and $S(p)$ is a subjective probability that is associated with outcome, x. In prospect theory $S(p)$ is a *decision weight* rather than a probability estimate, per se. These decision weights are assumed to increase monotonically with objective probabilities of events, but are larger than the objective probabilities for extremely unlikely outcomes and smaller than the objective probabilities for more likely outcomes. In prospect theory the decision weights for complementary events with probabilities p and $1-p$ need not necessarily add to one, but will generally be less than one, a property that Kahneman and Tversky [13] label as *subcertainty*.

However, since both SEU and prospect theory are based on the same simple bisymmetric model, they do not allow for a *differential* weighting of an event's probability in winning versus losing or "good" versus "bad" contexts as we have recently found [9, 15, 16]. That is, the models do not allow for the possibility that a decision maker might weight or even evaluate a probability like .2 or .8 differently, depending on the outcome with which it is associated. To account for such findings one needs a modification of SEU with a *dual* probability function. Such a model has been formally proposed by Luce and Narens [14]. Their dual bilinear model would allow for the measurement of probability bias, where "good" and "bad" outcomes can differentially affect subjective judgments of the same explicitly stated probabilities.

THE QUANTITATIVE METHOD

A modification of the procedure used originally by Davidson, Suppes, and Siegel [3] is now a computer-based algorithm, DSS, (cf. [16]) that independently measures the utility and subjective probability functions (U and S) in Eq. 1 above. Specifically, the S function is measured separately as two functions, S^+, and S^-, in order to assess potential bias in judgments of the likelihoods of good (S^+) and bad (S^-) outcomes. To the extent that the same function is obtained for S^+, and S^-, no context bias of this type is present in a decision maker's probability estimates. To the extent that the functions obtained for S^+, and S^- differ, a *measurable* cognitive bias exists in the individual's probability estimation process.

The DSS procedure involves determining the equality or indifference point in sequences of pairs of gambles so that Eq. 1 can be found for each gamble, and these equations can then be equated and solved in order to estimate the subjective utility associated with various outcomes. This utility function is then used in a second phase to estimate the subjective probability functions. On each trial, an individual is presented with two two-outcome gambles and is asked to indicate which of the two s/he prefers. The two-outcome gambles are set up as follows: Individuals are told that in each gamble, one outcome would be obtained if the event *E* occurs and the other outcome would be obtained if the event *E* does not occur. The event *E* is never specified, but individuals are informed that it has a true computer-generated probability of one-half. (Data from several studies have indicated that such instructions produce no bias between these two alternatives of *E* and ***not E***; individuals indeed weight the two events equally.) On each trial, one gamble, Gamble 1, has both outcomes fixed at specified values (e.g., some amount of money or points); the other gamble, Gamble 2, has one fixed outcome and one that is varied. For each of a specified number of trials (eight in the current version of DSS), the individual is asked to compare Gamble 1 with Gamble 2. The variable outcome in Gamble 2 is modified by DSS contingent upon the individual's response.

The decision maker's task on each trial is simply to indicate which gamble s/he prefers. If Gamble 1 is preferred, the variable outcome in Gamble 2 is adjusted upward by DSS to make this gamble more attractive; if Gamble 2 is preferred, the variable outcome in Gamble 2 is lowered to make this gamble less attractive. The amount of adjustment made by DSS depends on whether the individual indicates that one gamble is either slightly or strongly preferred to the other. Since events *E* and ***not E*** have probabilities fixed at .5 and these events are weighted as equivalent in probability, DSS determines the subjective utility function by noting in the variable-outcome gamble the value/amount necessary for a subject to change his/her preference ordering between the fixed-outcome and variable-outcome gambles (indicating the indifference or equivalence point for that pair of gambles). That is, DSS notes the amount that the individual assigns to the variable outcome in Gamble 2 such that s/he no longer has a clear preference for either Gambles 1 or 2.

One sequence of pairs of gambles we have used with DSS is presented in Table I; we will use these values throughout the remainder of this paper as an illustration. Each of the eight situations presented in Table I actually consists, then, of a series of adjustments to Gamble 2 that lead to the estimation of a subjective utility scale. For example, for the sequence presented in Table I, in the first situation, the individual is faced with one gamble for which s/he would lose \$10 with $p = .5$ (i.e., if *E* occurs) and would lose \$10 with $p = .5$ (if ***not E*** occurs). This, then, is a sure-loss gamble. The alternative gamble in the pair is described as resulting in a loss of **-\$A** dollars with $p = .5$, and a gain of \$10 with $p = .5$. The money amount associated with **-\$A** is initially randomly set to a large negative value or to a large positive value making one gamble initially more attractive. The individual adjusts the variable value up or down as necessary to reach indifference between the gambles (-\$10, -\$10) and (**-\$A**, +\$10). The final dollar amount associated with **-\$A** is recorded for the individual by DSS so that this information can be used to determine other utility values in subsequent Trials 3, 4, and 6.

Table I
Construction Sequence for Trials 1 - 8 Used to Find Individual Utility Functions

Trl	Gamble 1 Get	Gamble 1 Get	Gamble 2 Get	Gamble 2 Get	Subjective Utility
1	-10	-10	- ***A***	+10	***V(-A)*** = - 3
2	+10	+10	-10	+ ***B***	***V(+B)*** = + 3
3	-10	+ B	- A	+ ***C***	***V(+C)*** = + 5
4	- A	+10	- ***D***	+ B	***V(-D)*** = - 5
5	- D	+ B	- ***E***	+ C	***V(-E)*** = - 7
6	- A	+ C	- D	+ ***F***	***V(+F)*** = + 7
7	- D	+ F	- E	+ ***G***	***V(+G)*** = + 9
8	- E	+ C	- ***H***	+ F	***V(-H)*** = - 9

To facilitate the estimation process, we first assign a utility of +1 to + \$10 and a utility of -1 to -\$10. (Because in SEU theory the subjective utility scale is unique up to an affine transformation, that is, it is interval, we can without any loss of generality assign the utilities of +1 and -1.) Then, after **-\$A** is found, we can determine that the utility value, $V(-\$A)$, equals -3, by substituting in the formula in Eq. 1

$$S^{+}(.5) * (-1) + S^{-}(.5) * (-1) = S^{+}(.5) * (+1) + S^{-}(.5) * V(-\$A). \quad (2)$$

In a manner comparable to that for Trial 1 in Table I, a value for **+\$B** can be found next by comparing the gambles (+

$10, + $10) and (-$10, +$B), yielding the amount that is associated with a utility, V(+$B), of +3. These values of -$A and +$B are then used in Trials 3 - 8 to determine other points on the subjective utility scale. Currently, DSS finds an estimate for *+$C*, *-$D*, *-$E*, *+$F*, *+$G*, and *-$H*, which have utility value of +5, -5, -7, +7, +9, and -9, respectively as shown in Table I.

PROBABILITY ESTIMATION

Once these eight points on the utility function have been obtained along with -1 and +1, a nonlinear regression analysis is completed to find the best fitting utility curve for the estimated point values. A typical observed curve is shown in Figure 1. Given this best fitting curve, other utility values can then be estimated for the individual.

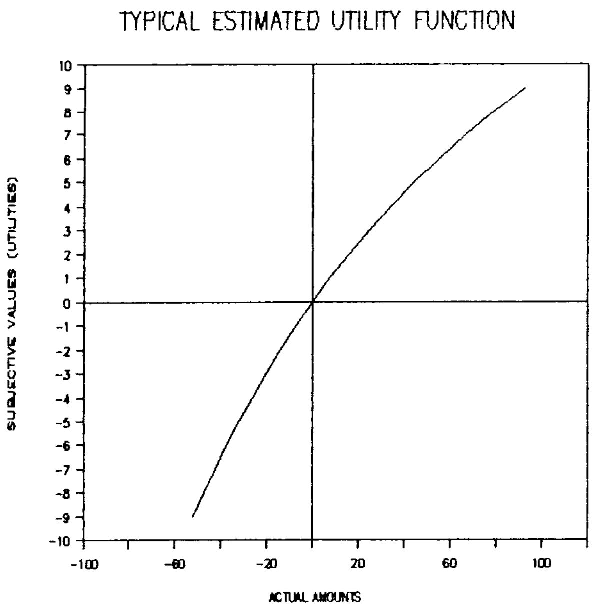

Figure 1. A Typical Utility Curve for Money. Losses End Is Steeper Than Gains End, Indicating Loss-Aversion.

A new series of gambles, now with different explicitly stated probability values, are used to obtain the subjective probability functions for winning (i.e., "good" outcomes) and losing (i.e., "bad" outcomes). Again, in the variable gamble, Gamble 2, one outcome is obtained if the event *E* (*S(p)* = .5) occurs and the other outcome is obtained if the event *E* does not occur (also *S(p)* = .5). The fixed gamble is similar to those presented in the utility estimation phase, except that now the probabilities of winning and losing are either .2/.8, .4/.6, .6/.4, or .8/.2. (Other values could also be programmed into DSS.) The variable outcome in Gamble 2 is again modified in a series of steps until the individual indicates that the two gambles are equally attractive. An example of such a gamble is shown in Figure 2.

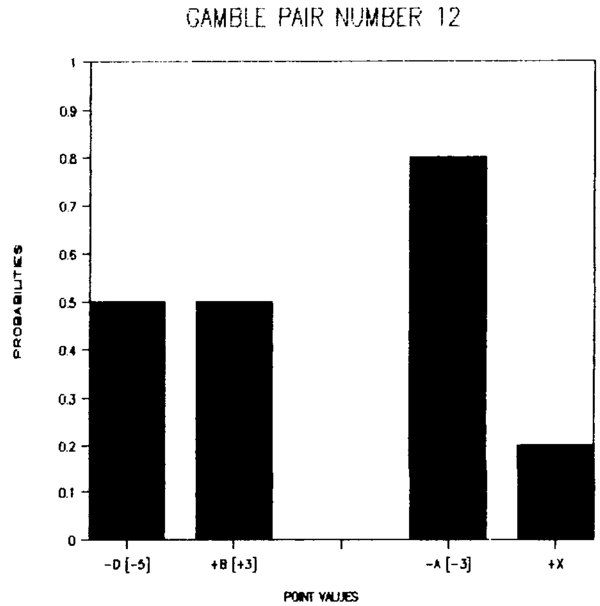

Figure 2. Example of Two Gambles in the Probability Estimation Phase.

Table II shows an illustrative series of eight trials used to estimate the probability functions. Letters **A - H** represent the amounts found previously on Trials 1 - 8, and X and Y are the variable outcomes; the stated probability values for each outcome are shown in parentheses. Note that on two pairs of trials (Nos. 9/11 and 13/15; Nos. 10/12 and 14/16) a subjective probability of winning value (S^+) and probability of losing value (S^-) is estimated for each of the values 2., .4, .6, and .8.

Table II
Construction Sequence for Trials 9 - 16 Used to Find Individual Probability Functions

	Gamble 1		Gamble 2	
Trl	Get	Get	Get	Get
9	- D(.5)	+ F(.5)	- A(.6)	+ X(.4)
10	- D(.5)	+ B(.5)	- H(.2)	+ X(.8)
11	- A(.5)	+ C(.5)	- D(.4)	+ X(.6)
12	- D(.5)	+ B(.5)	- A(.8)	+ X(.2)
13	- E(.5)	+ C(.5)	- D(.6)	+ Y(.4)
14	- A(.5)	+ C(.5)	- D(.2)	+ Y(.8)
15	- D(.5)	+ B(.5)	- E(.4)	+ Y(.6)
16	- A(.5)	+ C(.5)	-10(.8)	+ Y(.2)

ORIGINAL PAGE IS OF POOR QUALITY

Since events *E* and *not E* have probabilities fixed at .5 and the events are weighted by the individual as equivalent in probability, we can assign, with loss of generality, $S^+(.5) = S^-(.5) = .5$. Equations 3 and 4 illustrate two examples of trials where $S^+(.8)$ and $S^-(.2)$ are then determined as follows:

$$S^+(.5)*V(+\ \$B) + S^-(.5)*V(-\ \$D) = S^+(.8)*V(+\ \$X) + S^-(.2)*V(-\ \$H) \quad (3)$$

$$S^+(.5)*V(+\ \$C) + S^-(.5)*V(-\ \$A) = S^+(.8)*V(+\ \$Y) + S^-(.2)*V(-\ \$D) \quad (4)$$

If we assume that these two subjective probabilities act like objective probabilities and add to one, where $S^+(.8) = 1 - S^-(.2)$, then by substituting in utility values obtained from the estimated utility curve for -\$A, +\$B, +\$C, -\$D, -\$H, +\$X, and +\$Y, two independent estimates of $S^+(.8)$ and $S^-(.2)$ can be obtained by solving Eqs. 3 and 4. If we do not wish to assume that these two subjective probabilities add to one, we can still estimate $S^+(.8)$ and $S^-(.2)$ since we have two equations and two unknowns. In a comparable manner we can find restricted (i.e., add to one) or unrestricted (i.e., do not add to one) estimates of $S^-(.8)$ and $S^+(.2)$ from two new equations, and similarly for $S^+(.6)$ and $S^-(.4)$ and $S^-(.6)$ and $S^+(.4)$. Other probability values could also easily be estimated.

Regardless of whether a restricted or unrestricted model is fit, and regardless of the shape of each individual's subjective utility function, the S^+ and S^- functions should be the same if no probability bias due to a good vs. bad context exists in an individual's data. That is, if no bias exists, we should expect $S^-(.2) = S^+(.2)$, $S^-(.4) = S^+(.4)$, $S^-(.6) = S^+(.6)$, and $S^-(.8) = S^+(.8)$.

AN EMPIRICAL EXAMPLE

Subjects. Thirty-two male and female undergraduate students at The Ohio State University volunteered to participate in this study in partial fulfillment of the requirements of their introductory psychology course.

Procedure. On each trial, subjects were presented with two two-outcome gambles and were asked to indicate which of the two they preferred. In this way both subjective utilities and subjective probabilities for .2 - .8 were independently estimated. The subjects were instructed that they would be asked to make choices between pairs of gambles, and after they had indicated all of their preferences, some of the choice situations would be randomly sampled and played. The gamble that they would play in each pair would be randomly picked between the two gambles indicated to be equally attractive on each trial. Each subject was given 100 points, representing his or her credit for participating and was told that s/he would be gambling with this credit (not money) in the randomly selected gambles. Subjects were told that as a result of the gambling, they might either lose their credit hour if they lost their 100 points, retain their credit hour if they finished with more than zero points, or gain an additional credit if they finished with more than 200 points. At this point, all subjects were given the opportunity to withdraw from the study without penalty. None did so.

Results. Table III presents the estimated subjective probabilities that were obtained for .2, .4, .6, and .8 for both the winning and losing contexts. The data labeled "restricted" represents, as discussed above, the average of two estimates that are based on the restriction that the estimates for complementary events sum to one. The data labeled "unrestricted" are based on the estimates found when the sums are not restricted to add to one. Regardless of which of these models is assumed, the estimates in the table indicate that indeed across individuals a strong biasing effect exists. The same objectively stated probability values (.2, .4, .6, and .8) when presented to individuals, elicit consistently different subjective values or weights in their decision making process. At all four estimated levels of probability, the estimates that were associated with winning outcomes were consistently weighted lower than the corresponding values for losing. This represents a very strong affective bias in the estimation and/or weighting processes for subjective likelihoods.

Table III
Median Estimates of Winning and Losing Subjective Probabilities

	Restricted Case		
Actual	**Prob L**	**Prob W**	**Difference**
.200	.347	.208	.139
.400	.462	.401	.061
.600	.599	.539	.060
.800	.793	.653	.040

	Unrestricted Case		
Actual	**Prob L**	**Prob W**	**Difference**
.200	.312	.238	.074
.400	.469	.400	.069
.600	.567	.508	.059
.800	.815	.514	.301

SUMMARY

When an event has an established probability associated with it, it should be irrelevant whether that event is associated with a "good" or "bad" outcome context; the affective nature of the outcomes should not influence probability estimation or weighting in the decision process. In this study, two probability functions were estimated, and different values were found to be assigned to the winning and losing events. This differential weighting of the same event suggests that there is an affective influence on probability estimation in the decision making process. Individuals make choices between alternatives by assigning *different* subjective probabilities or weights to the *same* explicit event depending on whether it has a positive affective component (winning) or a negative affective component (losing).

The DSS methodology is important because it has the potential of not only quantitatively measuring this probability bias but also of explaining how some biases in probability estimation may cause suboptimal decisions, and how such bias can be reduced or eliminated in training decision makers to become more "expert" judges. It is designed to lead to programmatic research that has the ultimate application of developing training procedures that can: (a) standardize probability estimation methods in decision making under risk, (b) eliminate estimation biases such as over- and underestimation, (c) reduce individual differences in probability estimation, and (d) develop a scale for assessing a decision maker's accuracy and unbiasedness in subjective probability estimation. Research in this area is necessary if we are to go beyond merely describing suboptimal decision making behavior. The present study is an attempt to begin programmatic research that will allow us to predict suboptimal behavior due to biases in probability estimates and to train individuals to reduce bias in their judgments. In particular, several issues seems to be initially relevant to continued research. First, how strong and how generalizable is the differential weighting effect for probabilities? Second, what factors influence the strength of this effect? And, third, can a method like the DSS procedure coupled with the dual bilinear model allow us to predict and quantitatively measure the effects of suboptimal decision making strategies?

The research reported here, together with that reported elsewhere [8, 9, 10, 15, 16] is resulting in a more complete picture of the complex role that estimation bias and affect play in decision-making under risk or uncertainty. Our findings suggest that the dual bilinear model is a model worth pursuing. It has the potential to explain a number of difficult findings in the decision making literature including the framing effect [13], the differential weighting effect [16], and the "cautious optimists" effect [15]. Finally, the DSS procedure offers a quantitative measurement procedure for actually measuring rather that simply describing biases in judgment and decision making processes.

ACKNOWLEDGMENTS

This research was supported in part by Contract F33615-85-D-0514 from the U. S. Air Force and Grant S-760-7MG-052 from the Air Force Office of Scientific Research, Bolling AFB, DC to the first author.

REFERENCES

1. Arkes, H. R., and Hammond, K. R., (Eds.), *Judgment and decision making: An interdisciplinary reader*, (Cambridge University Press, New York, 1986).

2. Bar-Hill, M., "The base rate fallacy in probability judgments," *Acta Psychologica*, *44*, 1980, 211-233.

3. Davidson, D., Suppes, P., & Siegel, S., *Decision making: An experimental approach*, (Stanford University Press, Stanford, CA, 1957).

4. Edwards, W., "Conservatism in human information processing," in: B. Kleinmuntz (Ed.), *Formal representation of human judgment*, (Wiley, New York,1968), 17-52.

5. Einhorn, H. J., & Hogarth, R. M., "Behavioral decision theory: Processes of judgment and choice," *Annual Review of Psychology*, *32*, 1981, 53-88.

6. Fischoff, B. "Perceived informativeness of facts", *Journal of Experimental Psychology : Human Perception and Performance*, *3*, 1977, 349-358.

7. Fischoff, B., & Slovic, P, "A little learning ...: Confidence in multicue judgment tasks," in: R. E. Nickerson (Ed.), *Attention and performance (Vol. 8)*, (Erlbaum, Hillsdale, NJ, 1980), pp. 779-800.

ORIGINAL PAGE IS
OF POOR QUALITY

8. Hogarth, R., "Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics," *Psychological Bulletin, 90*, 1981, 197-217.

9. Isen, A. M., & Nygren T. E., "The influence of positive affect on the subjective utility of gains and losses: It's just not worth the risk," *Journal of Personality and Social Psychology*, in press, 1988.

10. Kahneman, D., Slovic, P., & Tversky, A., (Eds.), *Judgment under uncertainty: Heuristic and biases*, (Cambridge University Press, New York, 1982).

11. Kahneman, D., & Tversky, A., "Subjective probability: A judgment of representativeness," *Cognitive Psychology, 3*, 1972, 430-454.

12. Kahneman, D., & Tversky, A., "On the psychology of prediction," *Psychological Review, 80*, 1973, 237-251.

13. Kahneman, D., & Tversky, A., "Prospect theory: An analysis of decisions under risk," *Econometrika, 47*, 1979, 263-291.

14. Luce, R. D., & Narens, L., "Classification of concatenation measurement: Structures according to scale type," *Journal of Mathematical Psychology, 29*, 1986, 1-72.

15. Nygren, T. E., & Isen, A. M., "Examining probability estimation: Evidence for dual subjective probability functions," Paper presented at the Psychonomic Society Meetings, Boston, MA, November, 1985.

16. Nygren, T. E., & Morera, O., "Davidson, Suppes, & Siegel revisited: Evidence for a dual bilinear model," Paper presented at the Mathematical Psychology Meeting, Evanston, IL, July, 1988.

17. Slovic, P., Fischoff, B., & Lichtenstein, S., "Facts and fears: Understanding perceived risk," in: R. Schwing & W. A. Albers, Jr. (Eds.), *Societal risk assessment: How safe is safe enough?* (Plenum Press, New York, 1980) Plenum., 181-216.

18. Slovic, P., Fischoff, B., & Lichtenstein, S., "Response mode, framing, and information-processing effects in risk assessment," in: R. Hogarth (Ed.), *New directions for methodology of social and behavioral science: No.11, Question framing and response consistency*, (Jossey-Bass, San Francisco), 1982, 21-36.

19. Tversky, A., & Kahneman, D., "Availability: A heuristic for judging frequency and probability," *Cognitive Psychology, 5*, 1973, 207-232.

20. Tversky, A., & Kahneman, D., "Judgment under uncertainty: Heuristics and biases," *Science, 185*, 1974, 1124-1131.

21. Wickens, C. D., Stokes, A., Barnett, B., & Davis, T. "Componential analysis of pilot decision making," USAF Technical Report, June 1987.