1989

The effect of characteristics of the data distribution on subjective probability judgments.

Stephen H. Kleiman

University of Windsor

THE EFFECT OF CHARACTERISTICS OF THE DATA DISTRIBUTION ON SUBJECTIVE PROBABILITY JUDGMENTS

by
Stephen H. Kleiman
B.Sc. (Psychology), University of Michigan, 1980
M.A. (Psychology), University of Windsor, 1983

A Dissertation
Submitted to the
Faculty of Graduate Studies and Research
Through the Department of Psychology
in Partial Fulfillment
of the Requirements for the Degree
of Doctor of Philosophy at
the University of Windsor

Windsor, Ontario, Canada
1989

ISBN 0-315-61028-X

ABSTRACT

Accuracy in making probability judgments about the characteristics of data distributions was examined in this study. In particular, the effects of four characteristics on probability estimation: problem type, variability of the data distribution, amount of information, and anchor point were investigated. Previous research has examined the impact of single factors on single measures of subjective probability (usually the mean), but has not studied the interaction of factors or used more than one dependent measure. A questionnaire containing two problem scenarios was given to 747 introductory psychology students. One scenario dealt with the amount of rainfall in cm. per year in Buenos Aires, Argentina, and the other scenario was concerned with the average number of points scored per game in one season by a pro basketball team. Both problem scenarios were fictitious but appeared genuine to the subjects. A list of numbers representing data from preceding years followed each problem scenario. Each list was derived from computer generated random normal distributions and had either low or high variability, five or 25 years of data, and a last data point of 90, 100, or 110 units. Subjects were asked to generate a three category subjective probability distribution for the present year for each scenario. They also indicated their level of confidence in these probability judgments on a six point scale. Subjects also made point estimates of the mean from memory for each scenario. Four dependent measures resulted: the mean and standard deviation of the subjective probability distribution, the confidence judgment, and the point estimate of the mean from memory. The major findings of the study were:

1. The mean of the subjective probability distribution was affected by the interaction of anchor point and amount of information. The mean estimate was biased in the direction of the anchor point in the low amount of information condition but not in the high amount of information condition.

2. The standard deviation of the subjective probability distribution was affected by problem variability. Standard deviations were larger for the high variability condition than for the low variability condition. Also, females' judgments were more variable than males' judgments.

3. Confidence judgments were lower for the high variability condition than for the low variability condition. Confidence judgments were also affected by gender and by a gender by problem type interaction. Overall, males were more confident about their judgments than females. Further, males were more confident about the basketball problem than about the rainfall problem while females were equally confident about both problems.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

CHAPTER I

Introduction

During the 1950s and 1960s, normative decision theory was the predominant conceptual framework in the psychology of decision making (Wallstein, 1980). This theory assumed that individuals could select correct decision strategies and could make optimal choices (Edwards, 1954, 196l). However, research in the 1970s and 1980s revealed that individuals sometimes depart from the behavior predicted by normative decision theory (Kahneman & Tversky, 1984). Several studies showed that people make a number of judgmental errors in certain contexts. These errors include failure to understand regression, ignorance of sample size and base rate, and incorrect sample selection (Einhorn & Hogarth, 1978; Nisbett & Ross, 1980).

Some researchers suggest that even though the identification and description of these judgmental errors is valuable, it does not completely portray the capabilities of decision makers. They maintain that it is equally valuable to identify the judgmental strengths of decision makers. For example, other studies have found that people can use task information to correctly estimate descriptive statistics such as means or proportions (Peterson & Beach, 1967). In addition, some psychologists have suggested that previous research has overstated the judgmental shortcomings of decision makers (Christensen-Szalanski & Beach, 1984). The present study will examine whether people make normatively appropriate judgments about the characteristics of data distributions in the context of decision making under uncertainty.

Theoretical Framework

Decision making under uncertainty. Decision making under uncertainty has been an important framework for decision research. In the context of probability estimation, uncertainty refers to decisions to which no definite probability can be assigned. Although the probability of a given event is between zero and one, it can be difficult to determine what the exact probability is or to specify rules for determining that probability. Research has revealed that individuals often make incorrect estimates of uncertain events (Kahneman, Slovic & Tversky, 1982). Estimation errors can occur due to both task factors such as task complexity and ambiguity; and subject factors such as insufficient knowledge (Newell & Simon, 1972; Nisbett & Ross, 1980). The present study examines decision making under uncertainty by manipulating various distributional characteristics and observing how subjects' responses are affected.

In a classic article, Howell & Burnett (1979) developed a cognitive taxonomy for perceived probability or uncertainty. They suggested that people formulate an intuitive analog of the statistical properties of events and use these uncertainty estimates in combination with other information as a basis for decision making. The present study can be placed in the theoretical framework suggested by Howell & Burnett for decision making under uncertainty. First, it uses a probability estimation task that is frequentistic; that is, it uses categorical probability judgments (Howell & Burnett, 1979). Second; the decision task has two "generators": a known external generator (data lists) and an unpracticed internal generator (the cognitive processes used in decision making). For the known external generator, the cognitive elements that are stressed are prior generator knowledge and selective bias. Finally, the cognitive elements that are stressed for the unpracticed internal generator are frequency records and confidence bias (Howell & Burnett, 1979). Howell & Burnett's taxonomy is particularly useful for describing the present study because it demonstrates that this study is derived from past research on decision making under uncertainty.

Subjective probability judgments. Subjective probability judgments have been and remain an important dependent measure in decision research (Edwards, 196l). They represent an individual's estimate of the likelihood of a given event. This estimate may or may not correspond to the objective probability. In this study, subjective probability judgments are used as a dependent measure of subject responses. Careful inspection of these judgments should indicate whether subjects made normatively appropriate responses. Used in this manner, subjective probability judgments can be a valuable research tool.

Psychologists have studied subjective probability estimation in a variety of contexts, including: probability learning, intuitive statistics and risk (Cohen, Dearnaley, & Hansel, 1957; Peterson & Beach, 1967). Cohen, Hansel and their associates have discovered that age, prior experience, and the number and value of the alternatives offered can greatly affect subjective probability judgments. They have also explored various facets of the relationship between subjective probability and objective probability (Cohen, Dearnaley, & Hansel, 1956; Cohen & Hansel, 1955). Based on the results of this research, Cohen &

Hansel (1955) concluded that the relationship between subjective and objective probability is complex -- coinciding in some situations and systematically differing in other situations.

Coombs, Dawes and Tversky (1970) reported that individuals often produce subjective probability estimates that systematically depart from the appropriate objective probability values. In a study of subjective probability judgments, Kleiman (1983) reported that subjects estimated binomial distributions with 10 cases better than binomial distributions with 100 or 1000 cases. Analysis of the subjective probability estimates indicated that the subjects utilized the probability level of a distribution but ignored its sample size. Subjects were probably not aware that increasing sample size affects the standard error of a binomial distribution and; as a result, made serious judgmental errors.

Careful examination of the subjective probability estimates revealed that subjects underestimated the middle categories which have high objective probability values and overestimated categories at the tails, which have low objective probability values. Because subjects were unable to determine which distributional attributes were important in making their judgments, they did not correctly estimate the various categorical probabilities. As a result, they saw binomial distributions as more variable than they actually are when N is large.

The research cited above demonstrates that previous researchers have determined that subjective probability judgments are a valuable tool in studying decision making. In relationship to the present study, subjective probability judgments will be used to calculate the subjects' mean and standard deviation of their categorical probability judgments. These values will be used to determine if individual distributional characteristics (e.g., variability) affected subjects' probability judgments and if the subjects made normatively appropriate responses.

Estimation of the characteristics of data distributions.

Data distributions can be characterized by a number of factors, including central tendency, variability, sample size, and the number and nature of outlying data points. Variability is one of the main characteristics of a distribution. Differences in variability affect probability judgments because individuals can detect relative differences in variability (Peterson & Beach, 1967). Studies have also shown that distributions with large variances tend to be underestimated and distributions with small variances tend to be overestimated (Kahneman & Tversky, 1972; Starr & Kleiman, 1985). In the present study, variability was manipulated by using data lists with small and large standard deviations.

Studies have shown that changes in task information affect probability judgments (Payne, 1982; Wright, 1985). The amount of information or sample size is an important feature of any distribution. Samples based on small numbers of cases may not be representative of the parent population and may be unduly influenced by outlying values (DeVore, 1987). Moreover, people make different subjective probability judgments based on large amounts of information than on small amounts of information (Solso, 1979). In this study, the amount of information in a distribution will be manipulated by giving subjects distributions with different list lengths. Subjects will be given five years of data in the low information condition and 25 years of data in the high information condition.

In many decisions, people make estimates by adjusting an initial value upward or downward to yield a final answer. Subjects typically fail to accurately adjust the initial value (Tversky & Kahneman, 1974). Anchoring effects occur when subjects make estimates that are biased towards this initial value. In the present study, anchor point was defined as the last piece of information in a problem and represented the most recent item of information. This definition of anchor point differs from Tversky and Kahneman's definition because subjects can use all of the data when making their judgments. However, it is predicted that the last item of information will be used as a baseline for subjects' judgments. This study used three different values for the anchor point: 90 (below the problem mean), 100 (at the problem mean), and 110 (above the problem mean). The purpose of using various anchor points was to see if judgments would actually be biased in the direction of the last piece of information and to see how discrepant data points (i.e., outlying values) will affect subjects' judgments.

As discussed above, the present study examines the effects of three distributional characteristics (i.e., variability, sample size, and outlying data points) on subject responses. Different levels of these characteristics should affect subjects' probability judgments, confidence judgments, or point estimates. The intent of these manipulations is to determine whether individual distributional characteristics affect specific dependent measures.

Multivariate manipulation of distributional characteristics.

Besides examining how accurately people make probability judgments about a distribution by manipulating individual characteristics of a data distribution, this study will attempt to determine if these characteristics combine to affect judgment. Previous research has mainly examined the impact of single factors on single measures of subjective probability (usually the mean) (e.g., Bar Hillel, 1979; Tversky & Kahneman, 1971) but has sometimes looked at the impact of several independent variables on probability judgments (Pratt, 1982). However, the effects of several distributional characteristics have not been examined in the present research context. In this study, four independent variables will be simultaneously manipulated to determine if distributional characteristics will interact to affect probability judgments, as assessed by several dependent variables.

Problem type.

Task characteristics such as problem type can also affect subjective probability judgments (Slovic, Fischhoff, & Lichtenstein, 1976). Kahneman and Tversky (1984) reported that even minor changes in the wording of a problem can affect judgments. Therefore, the context and wording of a decision problem are important and can influence subject responses. In the present study, problem type will be manipulated by using two problem scenarios; a weather problem and a sports problem. Although the scenarios will appear genuine to the subjects, they will actually be fictitious. They were selected for the present study for two reasons; first, the rainfall scenario served as a gender neutral problem and the basketball scenario served as a male specific problem. Second, the scenarios permitted the same data distributions to be used in both problems,

Accuracy in estimation of meteorological events.

Meteorologists have to make probability judgments about rainfall amounts and temperature on a daily basis. Because the present study uses a weather problem, it might be worthwhile to review previous research about the accuracy of meteorological forecasting. Forecasters from the National Weather Service have been making probability estimates for precipitation and temperature since 1965. Studies of their computer assisted forecasts revealed that they consistently make high quality probability judgments for both precipitation and temperature (Murphy & Winkler, 1974; 1977). Murphy and Winkler (1974) found that their interval temperature forecasts deviated by only .028 from perfect calibration. The meteorologists' performance was equally impressive on precipitation forecasts, deviating only slightly from perfect calibration (Murphy & Winkler, 1977). Murphy (cited in Wright, 1985) reports that the data for recent years are even better. He attributed their accuracy to the experience that they have gained over the years with probability judgments. Not only do the meteorologists make precipitation and temperature forecasts every day but they also receive feedback about the accuracy of the previous day's forecast.

The research cited above indicates that people can make accurate probability judgments given adequate experience and feedback. The present study examines whether undergraduate college students can also make normatively appropriate judgments about a weather problem. Even though several studies have documented the judgmental errors that people make (e.g., Tversky & Kahneman, 1974; Hogarth, 198l), it is proposed that college students can make normatively appropriate probability judgments for data distributions given sufficient information.

Overconfidence in subjective probability estimation.

A number of studies on subjective probability estimation have found that subjects tend to be overconfident in their judgments (Slovic, Fischhoff, & Lichtenstein, 1977; Pitz & Sachs, 1984). Although subjective probability judgments tend to be positively related to the relative frequencies with which events occur, decision makers have displayed a systematic overconfidence bias in their responses (Lichtenstein, Fischhoff, & Phillips, 1977). If the events judged are discrete, they assign high subjective probabilities to events that actually occur less often and low subjective probabilities to events that actually occur more often. Similarly, if the events judged are continuous, individuals make judgments that are too tightly clustered around the mean of the subjective probability distribution, thereby ignoring events occurring in the tails. Lichtenstein, Fischhoff, & Phillips (1982) concluded that people are more likely to be overconfident about difficult judgments than about easy judgments. In fact, the most difficult judgments produce the most overconfident responses.

Seaver, von Winterfeldt & Edwards (1978) argued that overconfident judgments resulted from basic response biases. They compared several procedures for eliciting subjective probability distributions for continuous variables and found that asking for probability distributions on the basis of category intervals generates more overconfidence than direct probability assessment methods (e.g., point estimation). Alpert and Raiffa (1982) reported similar results. Although people have the tendency to make overconfident judgments for difficult tasks, Lichtenstein, Fischhoff, and Phillips (1982) found that training in and experience with probability estimation can improve performance.

While the research discussed above defined overconfidence in terms of the pattern of subjective probability judgments, the present study uses a more direct measure of overconfidence: confidence judgments based on a six point rating scale. As a result, subjects reported how confident they were about their subjective probability judgments.

Relationship between variability, uncertainty, and confidence. One objective of this study is to examine the relationship between variability, uncertainty and confidence judgments. Variability and uncertainty are closely related but not synonymous concepts (Tversky & Kahneman, 1982). Kahneman and Tversky (1982) found that uncertainty judgments were based on the distributional characteristics of data. Yates and Zukowski (1976) reported that uncertainty increased as the variability of the subjective probability distributions increased. Taken together, these findings indicate that high variability is associated with high uncertainty.

Moreover, Slovic, Fischhoff, and Lichtenstein (1980) found that high levels of uncertainty are associated with lower confidence judgments. Confidence judgments represent an individual's uncertainty about a prediction, estimate or inference that has already been made (Kahneman & Tversky, 1982), and generally refer to the individual's confidence about his/her estimates. On the basis of previous research, it seems likely that high variability should be associated with lower confidence judgments. The present study should help to ascertain under what conditions this relationship is valid.

Gender.

Although past research about the effects of gender on decision making has not produced consistent results (Lichtenstein & Fischhoff, 1981; MacCoby & Jacklin, 1974), gender differences have been found in cognitive style (Hyde & Rosenberg, 1980), in numerical reasoning and in mathematical ability (MacCoby & Jacklin, 1974). For example, previous research has revealed that women are less confident than men about their ability to perform mathematical or spatial tasks, especially those that are male specific (Hyde & Rosenberg, 1980). The present study involves a mathematical task using two scenarios; the rainfall scenario which is gender neutral and the basketball scenario which is male specific. On the basis of the research discussed above, male and female subjects will probably produce different probability and confidence judgments for the rainfall and basketball scenarios in the present study.

Statement of Hypotheses

Variability is recognized as an important distinguishing characteristic of a data distribution (Devore, 1987). Peterson and Beach (1967) reported that subjects can detect relative differences in the variability of distributions presented to them. This study investigates the effects of data variability on subjective probability judgments by presenting data sets with varying standard deviations. Each subject will receive one problem with a small standard deviation and one problem with a large standard deviation. Differences in the size of the standard deviation presented should result in differences in the subjective probability distribution generated.

Hypothesis 1: Subjects will be able to detect relative differences in variability of the data distributions presented to them. Therefore, their subjective probability distributions for problems consisting of data with high variability will have larger judged standard deviations than problems consisting of data with low variability.

Research suggests that variability, uncertainty and confidence judgments are interrelated. Yates and Zukowski (1976) found that high variability is associated with high uncertainty . Slovic, Fischhoff, and Lichtenstein (1980) also reported that high levels of uncertainty about a problem are reflected in lower beliefs about judgmental accuracy. On the basis of this research and using uncertainty as a mediating variable, it appears that high variability is associated with lower confidence judgments. In the present study, there are both low and high variability conditions. Differences in variability should result in differences in the subjects' confidence judgments.

<u>Hypothesis 2</u>: Variability will affect confidence judgments. Subjects will give problems with large standard deviations lower confidence judgments than problems with small standard deviations.

The amount of information or sample size is recognized as an important distributional characteristic (DeVore, 1987). Samples based on large amounts of information are more likely to be representative of their parent populations than samples based on small amounts of information. Because small sample sizes are associated with larger standard errors than large sample sizes, subjects should have higher levels of uncertainty about distributions based on five years of data than distributions based on 25 years of data. Slovic, Fischhoff, and Lichtenstein (1980) reported that high levels of uncertainty about a problem are reflected in lower beliefs about judgmental accuracy. Therefore, differences in the amount of data in a problem should result in differences in the subjects' confidence judgments.

<u>Hypothesis 3</u>. Amount of information will affect confidence judgments. Subjects will give problems based on 25 years of data higher confidence judgments than problems based on five years of data.

Although it is predicted that amount of information will not affect subjects' probability judgments by itself, decision research suggests that the interaction of amount of information and anchor point should affect their subjective probability judgments (Slovic & Lichtenstein, 1971; Kahneman & Tversky, 1972). The present study has both low and high information conditions and uses three different anchor points (90, 100 and 110). It follows from normative statistical principles that extreme values should affect mean estimates more when the amount of data is small than when it is large. If the subjects make normatively appropriate judgments, then discrepant anchors (90 or 110) will affect mean estimates more in the low information condition than in the high information condition.

Hypothesis 4: The amount of information and the anchor point will interact to affect the mean categorical probability estimate (MCAT). Mean estimates will be biased in the direction of the anchor point for the low information condition but not for the high information condition.

Subjective probability judgments can also be affected by task characteristics such as problem type (Slovic, Fischhoff, & Lichtenstein, 1976). Kahneman and Tversky (1984) reported that even simple changes in the wording of a problem can affect judgments. Probability judgments are also affected by an individual's general knowledge of people, objects and events (Nisbett & Ross, 1980). In general, people are more confident about tasks and events that are familiar to them and that are gender appropriate (Lenney, 1977). Two problem scenarios; the rainfall problem and the basketball problem, are used in this study to examine how problem type affects confidence judgments. They were used because the rainfall scenario was constructed to be a gender neutral problem while the basketball scenario was constructed to be a male specific problem.

Hypothesis 5: Problem type and gender will interact to affect confidence judgments. Men will probably believe that they are more knowledgeable about basketball scores than about amounts of rainfall, and therefore will be more confident about their judgments for the basketball problem than for the rainfall problem. Women will probably believe that they are equally knowledgeable about both basketball scores and amounts of rainfall, and will have similar confidence judgments for both problems.

CHAPTER II

Method

Subjects

The subjects were 747 undergraduate students enrolled in the introductory psychology courses at the University of Windsor. They had a mean age of 20.69 years and a sex distribution of 276 males and 471 females. Each subject received one course credit point for participating in the study.

Procedure

With the permission of each instructor, the experimenter arrived at the classroom at a pre-arranged time. After being introduced to the students, the experimenter gave the following instructions:

Today you have the opportunity to participate in a research study. This study is concerned with probability estimation. The experimental questionnaire is six pages long and takes only 15 minutes to complete. Additionally, the questionnaire is anonymous. The only personal information that you will supply is your age and your sex.

Although participation in this study is voluntary, it will be to your advantage to participate. If you fill out the questionnaire, you will receive one point of extra credit toward your grade.

A short description of the study and complete directions are contained on the first page of the questionnaire. Be sure to read the problem carefully. Any comments that you might have about the questionnaire would be appreciated. These comments may be written directly on the questionnaire. Now please raise your hand if you would like to fill out the questionnaire. Once you receive the questionnaire you may begin.

Subjects were then given a copy of the questionnaire. They had about 15 minutes to complete the questionnaire. Any questions that arose were answered. Each class tested received all of the different versions of the questionnaire.

Research Instrument

The experimental questionnaire contained six pages. The first page had a short description of the study and gave directions to the subjects. It also asked the subjects for demographic information. Pages 2-3 contained the first problem and pages 4-5 the second problem. Finally, the sixth page asked the subjects to give their point estimates for both problems. For each task, subjects gave subjective probability estimates, confidence judgments and point estimates. All versions of the experimental questionnaire are given in Appendix A.

All subjects received both problems: one rainfall problem and one basketball problem. Half the subjects received the rainfall problem first and the basketball problem second while the other subjects received the basketball problem first and the rainfall problem second. Each problem was constructed so that it contained one level of the three independent variables; i.e., each problem had one level of variability, amount of information, and anchor point. Each version of the questionnaire was constructed so that the problems contained different levels of each independent variable.

Levels of each independent variable were individually paired in the two problems. For example, high variability in the rainfall problem was paired with low variability in the basketball problem; high amount of information in the rainfall problem was paired with low amount of information in the basketball problem; and an anchor point of 110 in the rainfall problem was paired with an anchor point of 90 or 100 in the basketball problem. Therefore,no subject received the same level of any independent variable in both problems.

Twenty-four versions of the experimental questionnaire were constructed. Variability had two levels, amount of information had two levels, and anchor point had three levels; and each problem appeared in the first position in twelve versions and in the last position in twelve versions. The 2 X 2 X 3 (X 2 problems) experimental design used is given in Table 1. Samples of the two problems are given in Table 2.1 and 2.2. Abbreviations used in the text and in all subsequent tables are defined in Appendix B.

Construction of the tasks. Two scenarios were created for the study. The amount of annual rainfall in Buenos Aires, Argentina was used in one scenario and the average points per basketball game scored by the Boston Celtics was used in the other scenario. They were designed to appear authentic to the subjects, but they were actually fictitious.

The data used in the problems were obtained from computer generated normal distributions with a sample size of 100 and a mean of 100. Two of the distributions generated had a standard deviation of two, and the other two distributions had a standard deviation of 10. One distribution for each of the standard deviations was selected for further use. For each of the distributions, lists containing five data points and 25 data points were constructed. Then the data in the lists were arranged so that they had similar patterns of high and low values.

After the four lists had been constructed, their means were calculated. The data were then adjusted so that the mean was 99.8 for each list. The last value in each list was 99.7. Finally, data points in the list were assigned sequential year markers, with the order of data points maintained. Although all lists had a mean of

TABLE 1

Experimental Design Layout by Variability (VAR) by Anchor Point (ANCHOR) by Amount of Information (INFO).

Low Amount of Information

	Anchor Point		
Variability	90	100	110
LOW	VAR = LOW INFO = LOW ANCHOR = 90 (61)	VAR = LOW INFO = LOW ANCHOR = 90 (59)	VAR = LOW INFO = LOW ANCHOR = 90 (62)
HIGH	VAR = HIGH INFO = LOW ANCHOR = 90 (63)	VAR = HIGH INFO = LOW ANCHOR = 100 (63)	VAR = HIGH INFO = LOW ANCHOR = 110 (65)

High Amount of Information

	Anchor Point		
Variability	90	100	110
LOW	VAR = LOW INFO = HIGH ANCHOR = 90 (62)	VAR = LOW INF0 = HIGH ANCHOR = 100 (64)	VAR = LOW INFO = HIGH ANCHOR = 110 (61)
HIGH	VAR = HIGH INFO = HIGH ANCHOR = 90 (62)	VAR = HIGH INFO = HIGH ANCHOR = 100 (63)	VAR = HIGH INFO = HIGH ANCHOR = 110 (65)

Note: Numbers in parentheses are the sample n's.

Table 2.1

Sample Rainfall Problem

VAR = HIGH

INFO = HIGH

ANCHOR = 110

Meteorologists in Buenos Aires, Argentina decided to examine the pattern of rainfall in Buenos Aires. The Archives of the Argentinian Meteorological Society showed that the measurable annual rainfall (in cm.) in Buenos Aires, Argentina for the years 1960 to 1984 was:

1960 : 110.0	1969 : 102.9	1977 : 99.1
1961 : 112.4	1970 : 85.0	1978 : 94.8
1962 : 92.9	1971 : 96.7	1979 : 108.8
1963 : 107.6	1972 : 90.4	1980 : 103.1
1964 : 114.8	1973 : 87.4	1981 : 96.2
1965 : 87.0	1974 : 100.3	1982 : 106.7
1966 : 93.4	1975 : 111.2	1983 : 108.1
1967 : 104.7	1976 : 86.7	1984 : 99.7
1968: 95.3		

In 1985, the annual rainfall in Buenos Aires was 110 cm.

Table 2.2

<u>Sample Basketball Problem</u>

VAR=LOW

INFO = LOW

ANCHOR = 90

A group of sportswriters in Boston, Massachusetts decided to examine the average points per game that the Boston Celtics scored over the past few seasons. <u>Pro Basketball Digest</u> showed that the average points per game scored by the Boston Celtics for the years 1980 to 1984 was:

1980 : 98.4

1981 : 99.1

1982 : 101.5

1983 : 100.4

1984 : 99.7

In 1985, the Boston Celtics scored an average of 90 points per game.

99.8, lists in the low variability condition were constructed to have smaller standard deviations than lists in the high variability condition.

In the low variability condition, lists with five data points had a standard deviation of 1.20 and lists with 25 data points had a standard deviation of 1.25. Lists in the high variability condition with five data points had a standard deviation of 6.00 and lists with 25 data points had a standard deviation of 8.65. The discrepancy in standard deviations occurred because data in the low information condition had to conform to certain requirements. First, the data had to be constructed so that extremely low or high values were not noticeable. Second, the pattern of high and low values in the low information had to be similar to the pattern used in the low variability condition.

Independent Variables

Four factors served as independent variables: 1) Problem Type 2) Variability 3) Amount of Information and 4) Anchor Point. Further gender served as a subject classification variable.

Problem type. Problem Type is defined as the type of problem scenario presented to the subjects. Two different problems were used in this study. Problem Type 1 dealt with the amount of annual rainfall in centimeters in Buenos Aires, Argentina and Problem Type 2 dealt with the average number of points per game scored by a pro basketball team.

Variability. Variability is defined as the dispersion of a data set. In this study, variability was manipulated by using standard deviations of different magnitudes. Some problems had data with low variability (standard deviation of approximately 1.20), while other problems had data with high variability (standard deviation of 6.00 or 8.65).

Amount of information. Amount of information is defined as the number of previous years of data given in a problem. List length was varied so that a list contained data for either five years (low information condition) or 25 years (high information condition). List lengths of five and 25 were selected because pretesting indicated that subjects gave different probability estimates and confidence judgments in the low information condition than in the high information condition. In addition, it was difficult for subjects in the high information condition to directly calculate the mean and the standard deviation during the experiment.

Anchor point. Anchor point is defined as the last data point (1985 data point) in a problem. In the present study, three anchor points were used: below the mean (90), at the mean (100), or above the mean (110).

Dependent Variables

Four dependent variables were used: The mean and the standard deviation of the subjective probability estimates, confidence judgments, and point estimates.

Subjective probability estimates. Subjects produced subjective probability estimates for each problem. An example of how the subjective probability estimates were produced is given below:

> Based on the information given on the preceding page, please estimate the probability that each of the following categories will occur (in %).

	Probability Estimate (in %)
In 1986, less than 90 cm. of rain will fall in Buenos Aires, Argentina	______
In 1986, 90 to 110 cm. of rain will fall in Buenos Aires, Argentina	______
In 1986, greater than 110 cm. of rain will fall in Buenos Aires, Argentina	______
	Total = 100%

Note: Your probability estimates should add up to about 100% because the categories include all possibilities.

The sum of each subject's subjective probabilities was 1.0. The expected value and the standard deviation of the subjective probability distribution (MCAT and SCAT respectively) were calculated for each subject and served as the dependent variables derived from the category judgments.

MCAT. The expected value of the subjective probabilitydistribution (MCAT) was calculated using a weighted average of the subjective probabilities.

MCAT = [(80 * CAT1) + (100 * CAT2) + (120 * CAT3)]

where CAT1, CAT2, and CAT3 are the subjective probability estimates for the three categories.

SCAT. The standard deviation of the subjective probability distribution (SCAT) was also calculated using a weighted average of the subjective probabilities; the formula was:

SCAT = [[((8O * 80) * CAT1) + ((100 * 100) * CAT2) + (12O * 120) * CAT3)] - [MCAT * MCAT)]] **1/2

where CAT1, CAT2, and CAT3 are the three subjective probabilities, and MCAT is the expected value for the categorical probability estimate.

Category means of 80 and 120 were used in the formulas to calculate MCAT and SCAT. Although different category means could have been selected, 80 and 120 were chosen because they fulfill certain criteria. Assuming that the subjective probability distribution is unimodal and symmetrical, category means of 80 and 120 permit category interval length to be preserved. Second, they partition the tail area of the distribution into two equal parts (i.e., the area between 0 and 80 is equivalent to the area between 80 and 90 and the area between 110 and 120 is equivalent to the area between 120 and ∞.) The actual category boundaries are 0 and ∞ .

Confidence judgments. Confidence judgments (CONJ) based on a six point scale were used as another dependent variable. The six point rating scale for the confidence judgments used adverb modifiers with scale values taken from empirical research (Cliff, 1959). These scale values placed the range of confidence judgments into equal intervals. Subjects indicated their confidence in the accuracy of their categorical probability estimates for both problems. An example of how confidence judgments were obtained is given below:

Please place an X at a point on the scale below which best describes your confidence in the accuracy of your probability estimates above.

________	________	________	________	________	________
Not At All Confident	Slightly Confident	Rather Confident	Quite Confident	Very Confident	Extremely Confident

Point estimates of the mean. Finally the subject's point estimates of the mean from memory (PTE) served as a dependent variable. Without looking back on the preceding pages of the questionnaire, subjects gave point estimates of the mean for both problems. An example of how the point estimates were obtained is given below:

> WITHOUT looking back on the previous pages of the questionnaire, please give your best estimate of the average amount of annual
> rainfall in cm. in Buenos Aires, Argentina:
>
> ____________ cm.

CHAPTER III

Results

The results for this study are reported in four sections:

1) Descriptive Statistics: Means and Ns for the rainfall problem, the basketball problem, and for both problems combined; and the correlation coefficients among the four dependent variables and age.

2) Major Analyses: Results of the 4 four-way ANOVAs for the effects of variability, amount of information, anchor point, problem type and/or their interactions on the four dependent variables (MCAT, SCAT, CONJ and PTE).

3) Secondary analyses: Results of the two-way ANOVA for the effects of gender, problem type and/or their interactions on the four dependent variables (MCAT, SCAT, CONJ and PTE).

4) Summary of results.

Descriptive Statistics

Tables 3.1-3.3 give the means and Ns for the four dependent variables (MCAT, SCAT, CONJ and PTE) for the rainfall problem; Tables 4.1-4.3 give the means and Ns for the four dependent variables for the basketball problem; and Tables 5.1-5.3 give the means and Ns for the four dependent variables for the rainfall and basketball problems combined. Table 6 gives the means and Ns for the four dependent variables broken down by gender and problem type.

TABLE 3.1

Means of the 4 Dependent Variables (MCAT, SCAT, CONJ, & PTE) for the Rainfall Problem in the Low Variability Condition (VAR) by Amount of Information (INFO) and Anchor Point (ANCHOR).

Rainfall Problem - Low Variability Condition

	Dependent Variables			
Anchor Point	MCAT	SCAT	CONJ	PTE
	Low Amount of Information			
90	90.26	8.60	3.31	95.67
100	100.7	7.23	3.63	98.59
110	101.7	8.37	3.58	99.45
Combined Anchor Points	100.0	8.08	3.51	97.91
	High Amount of Information			
90	99.63	7.89	3.55	96.87
100	99.96	7.94	3.64	98.88
110	100.4	7.50	3.52	98.93
Combined Anchor Points	100.0	7.78	3.57	98.23
	Combined Amount of Information			
90	98.95	8.24	3.43	96.28
100	100.3	7.60	3.63	98.74
110	100 .7	7.94	3.55	99.20
Combined Anchor Points	100.0	7.93	3.54	98.07

TABLE 3.2

Means of the 4 Dependent Variables (MCAT, SCAT, CONJ, & PTE) for the Rainfall Problem in the High Variability Condition (VAR) by Amount of Information (INFO) and Anchor Point (ANCHOR).

Rainfall Problem - High Variability Condition

Anchor Point	MCAT	SCAT	CONJ	PTE
		Dependent Variables		
		Low Amount of Information		
90	100.0	9.25	3.68	97.57
100	99.98	9.53	3.30	97.43
110	100.9	8.77	3.31	99.10
Combined Anchor Points	100.3	9.19	3.43	98.02
		High Amount of Information		
90	99.58	10.47	3.21	97.19
100	99.88	10.16	3.19	99.02
110	99.98	10.16	3.31	100.1
Combined Anchor Points	99.82	10.26	3.24	98.79
		Combined Amount of Information		
90	99.81	9.86	3.45	97.38
100	99.93	9.85	3.25	98.22
110	100 .4	9.49	3.31	99.61
Combined Anchor Points	100.1	9.73	3.33	98.41

TABLE 3.3

Means of the 4 Dependent Variables (MCAT, SCAT, CONJ, & PTE) for the Rainfall Problem for Both Variability Conditions Combined (VAR) by Amount of Information (INFO) and Anchor Point (ANCHOR).

Rainfall Problem - Both Variability Conditions Combined

	Dependent Variables			
Anchor Point	MCAT	SCAT	CONJ	PTE
		Low Amount of Information		
90	99.16	8.93	3.50	96.64
100	100.4	8.42	3.46	97.99
110	101.0	8.57	3.45	99.28
Combined Anchor Points	101.2	8.64	3.47	97.96
		High Amount of Information		
90	99.61	9.18	3.38	97.03
100	99.92	9.04	3.42	98.94
110	100.2	8.87	3.41	99.53
Combined Anchor Points	99.91	9.03	3.40	98.51
		Combined Amount of Information		
90	99.38	9.06	3.44	96.83
100	100.1	8.74	3.44	98.48
110	100 .6	8.72	3.43	99.41
Combined Anchor Points	100.0	8.84	3.44	98.24

TABLE 4.1

Means of the 4 Dependent Variables (MCAT, SCAT, CONJ, & PTE) for the Basketball Problem in the Low Variability Condition (VAR) by Amount of Information (INFO) and Anchor Point (ANCHOR).

Basketball Problem - Low Variability Condition

	Dependent Variables			
Anchor Point	MCAT	SCAT	CONJ	PTE
	Low Amount of Information			
90	99.04	9.00	3.40	96.94
100	100.0	7.43	3.65	98.27
110	101.8	9.09	3.52	99.48
Combined Anchor Points	100.3	8.51	3.52	98.21
	High Amount of Information			
90	99.49	8.56	3.56	96.98
100	100.0	8.62	3.37	99.85
110	101.2	7.96	3.95	99.75
Combined Anchor Points	100.2	8.38	3.63	98.86
	Combined Amount of Information			
90	99.26	8.79	3.48	96.46
100	100.0	8.02	3.51	99.06
110	101.6	8.53	3.74	99.61
Combined Anchor Points	100.3	8.45	3.58	98.54

TABLE 4.2

Means of the 4 Dependent Variables (MCAT, SCAT, CONJ, & PTE) for the Basketball Problem in the High Variability Condition (VAR) by Amount of Information (INFO) and Anchor Point (ANCHOR).

Basketball Problem - High Variability Condition

Anchor Point	MCAT	SCAT	CONJ	PTE
		Low Amount of Information		
90	99.89	8.82	3.42	95.88
100	100.2	8.53	3.38	96.98
110	101.3	9.43	3.38	98.57
Combined Anchor Points	100.1	8.93	3.39	97.17
		High Amount of Information		
90	101.1	10.11	3.23	96.77
100	101.0	9.76	3.65	99.98
110	99.72	9.76	3.40	99.44
Combined Anchor Points	100.3	9.87	3.43	98.41
		Combined Amount of Information		
90	99.52	9.46	3.33	96.33
100	100.6	9.14	3.51	98.47
110	100.5	9.60	3.39	98.50
Combined Anchor Points	100.2	9.40	3.41	97.79

TABLE 4.3

Means of the 4 Dependent Variables (MCAT, SCAT, CONJ, & PTE) for the Basketball Problem for Both Variability Conditions Combined(VAR) by Amount of Information (INFO) and Anchor Point (ANCHOR).

Basketball Problem - Both Variability Conditions Combined

Anchor Point	MCAT	SCAT	CONJ	PTE
	Dependent Variables			
	Low Amount of Information			
90	98.97	8.91	3.41	96.43
100	100.1	7.98	3.52	97.63
110	101.5	9.26	3.45	99.02
Combined Anchor Points	100.2	8.72	3.46	97.70
	High Amount of Information			
90	99.81	9.32	3.40	96.65
100	100.5	9.19	3.51	98.76
110	100.5	8.85	3.68	99.06
Combined Anchor Points	100.3	9.12	3.53	98.17
	Combined Amount of Information			
90	99.38	9.12	3.40	96.65
100	100.3	8.58	3.51	98.76
110	101.0	9.06	3.57	99.06
Combined Anchor Points	100.2	8.92	3.49	98.17

TABLE 5.1

Means of the 4 Dependent Variables (MCAT, SCAT, CONJ, & PTE) for the Basketball and Rainfall Problems Combined for the Low Variability Condition (VAR) by Amount of Information (INFO) and Anchor Point (ANCHOR).

Both Problems Combined - Low Variability Condition

	Dependent Variables			
Anchor Point	MCAT	SCAT	CONJ	PTE
		Low Amount of Information		
90	98.66	8.81	3.36	96.33
100	100.4	7.33	3.64	98.43
110	101.4	8.73	3.55	99.46
Combined Anchor Points	101.2	8.30	3.51	98.06
		High Amount of Information		
90	99.56	8.23	3.56	96.93
100	99.98	8.28	3.51	99.36
110	100.8	7.73	3.74	99.35
Combined Anchor Points	100.1	8.08	3.60	98.55
		Combined Amount of Information		
90	99.11	8.52	3.46	96.62
100	100.2	7.81	3.57	98.90
110	101.1	8.24	3.65	99.41
Combined Anchor Points	100.1	8.19	3.56	98.31

TABLE 5.2

Means of the 4 Dependent Variables (MCAT, SCAT, CONJ, & PTE) for the Basketball and Rainfall Problems Combined for the High Variability Condition (VAR) by Amount of Information (INFO) and Anchor Point (ANCHOR).

Both Problems Combined - High Variability Condition

	Dependent Variables			
Anchor Point	MCAT	SCAT	CONJ	PTE
	Low Amount of Information			
90	99.47	9.04	3.55	96.75
100	100.1	9.03	3.34	97.21
110	101.1	9.11	3.35	98.83
Combined Anchor Points	100.2	9.06	3.41	97.60
	High Amount of Information			
90	99.86	10.29	3.22	96.98
100	100.4	9.96	3.42	99.50
110	99.85	9.96	3.35	99.28
Combined Anchor Points	100.1	10.07	3.33	98.60
	Combined Amount of Information			
90	99.67	9.66	3.39	96.87
100	100.3	9.50	3.38	98.35
110	100.5	9.54	3.35	99.06
Combined Anchor Points	100.1	9.57	3.37	98.10

TABLE 5.3

Means of the 4 Dependent Variables (MCAT, SCAT, CONJ, & PTE) for the Basketball and Rainfall Problems Combined for Both Variability Conditions Combined (VAR) by Amount of Information (INFO) and Anchor Point (ANCHOR).

Both Problems Combined - Both Variability Conditions Combined

	Dependent Variables			
Anchor Point	MCAT	SCAT	CONJ	PTE
	Low Amount of Information			
90	99.06	8.92	3.45	96.53
100	100.2	8.20	3.49	97.81
110	101.3	8.92	3.45	99.15
Combined Anchor Points	101.2	8.68	3.46	97.83
	High Amount of Information			
90	99.71	9.25	3.39	96.96
100	100.2	9.12	3.46	99.43
110	100.3	8.86	3.55	99.31
Combined Anchor Points	100.1	9.08	3.47	98.58
	Combined Amount of Information			
90	99.38	9.09	3.42	96.74
100	100.2	8.86	3.47	98.62
110	100.8	8.89	3.50	99.23
Combined Anchor Points	100.1	8.88	3.47	98.20

TABLE 6

Means of the 4 Dependent Variables (MCAT, SCAT, CONJ, & PTE) by Gender (GENDER) and Problem Type (TYPE)

Gender	MCAT	SCAT	CONJ	PTE
		Problem Type — Dependent Variables		
		Rainfall		
Male	99.96	8.34	3.89	98.84
Female	100.1	9.13	3.17	97.89
Both Genders	100.0	8.84	3.44	98.24
		Basketball		
Male	100.6	8.14	4.22	99.26
Female`	100.0	9.38	3.07	97.52
Both Genders	100.2	8.92	3.49	98.17
		Both Problems Combined		
Male	100.3	8.24	4.06	99.05
Female	100.1	9.25	3.12	97.71
Both Genders	100.1	8.88	3.47	98.20

Note: There were 276 males and 471 females in the study

Table 7 gives the correlation coefficients among the four dependent variables and age as well as their means and their standard deviations. The correlations for age are given because it is the only continuous demographic variable. Although several *r* values are significant at $p < .01$, they may not be especially meaningful because the *r* values are at .170 or less. These results may have been obtained because the sample size was extremely large and because weak relationships may exist between the dependent variables. However this explanation can not account for the large correlation between SCAT and CONJ ($r = -.405$). This correlation indicates that the larger the standard deviation of the subjective probability distribution, the lower the confidence judgment.

Major Analyses

The effects of variability, amount of information, anchor point, and problem type. The ANOVA summary table for the four dependent variables (MCAT, SCAT, CONJ and PTE) by variability, amount of information, anchor point and problem type is given in Table 8.

A note about the use of error terms in the analysis of variance (ANOVA) summary tables is necessary. The number of respondents (747) precluded performing repeated measures across subjects. When attempts were made to do so, both CPU time and space proved insufficient for processing. Therefore, all ANOVA procedures were performed between rather than within subjects and one pooled error term was used. The degrees of freedom for the error term are large enough that the critical *F*-values would be almost identical to those for the repeated measures and the mean square error estimates are extremely stable. If anything, the *F*-tests would tend to be somewhat conservative.

Table 7

Correlation Coefficients and Descriptive Statistics for the

4 Dependent Measures (MCAT, SCAT, CONJ, and PTE) and Age

	MCAT	SCAT	CONJ	PTE	AGE
MCAT	----	-.016	.032	.154*	.045
SCAT		----	-.405*	-.110*	-.034
CONJ			—	.170*	-.066
PTE				—	-.069*
Mean	100.10	8.88	3.47	98.20	20.69
S.D.	3.91	3.83	1.29	6.42	4.67

*$p < .01$

Note: $N = 1494$

Table 8

ANOVA Summary Table for the 4 Dependent Variables (MCAT, SCAT,CONJ, and PTE) by Variability (VAR), Amount of Information (INFO), Anchor Point (ANCHOR), and Problem Type (TYPE).

		MCAT		SCAT		CONJ		PTE	
Source:	df	SS	F	SS	F	SS	F	SS	F
TYPE	1	14.41	.97	3.65	.26	1.22	.74	2.12	.05
VAR	1	.01	.00	710.00	50.49*	12.90	7.84*	18.10	.45
INFO	1	3.69	.25	60.40	4.30	.00	.00	201.00	5.02
ANCHOR	2	505.60	16.90*	47.20	1.68	1.57	.00	1678.00	20.90*
TYPE X VAR	1	1.21	.08	66.50	4.73	.13	.08	114.00	2.85
TYPE X INFO	1	9.75	.65	.03	.00	1.70	1.03	16.00	.40
TYPE X ANCHOR	2	9.59	.32	15.30	.54	1.80	.55	27.10	.34
VAR X INFO	1	1.33	.09	139.00	9.89*	2.55	1.55	24.70	.62
VAR X ANCHOR	2	90.80	3.05	19.30	.69	3.13	.95	41.20	.51
INFO X ANCHOR	2	160.60	5.39*	64.70	2.30	1.79	.54	150.00	1.86

Table 8 continued

		MCAT		SCAT		CONJ		PTE	
Source:	*df*	*SS*	*F*	*SS*	*F*	*SS*	*F*	*SS*	*F*
TYPE X VAR X INFO	1	12.24	.82	2.20	.16	.85	.51	.95	.02
TYPE X VAR X ANCHOR	2	56.99	1.91	7.66	.27	5.87	1.78	51.50	.64
TYPE X INFO X ANCHOR	2	20.25	.68	23.20	.83	.75	.23	42.50	.53
VAR X INFO X ANCHOR	2	38.40	1.29	72.50	2.58	8.47	2.57	46.50	.58
TYPE X VAR X INFO X ANCHOR	2	44.30	1.49	3.37	.12	5.30	1.61	73.00	.91
ERROR	1470	21898.00		20680.00		2421.00		59013.00	
TOTAL	1493	22865.00		21910.00		2470.00		61496.00	
R^2		.042		.056		.020		.040	

*$p < .01$

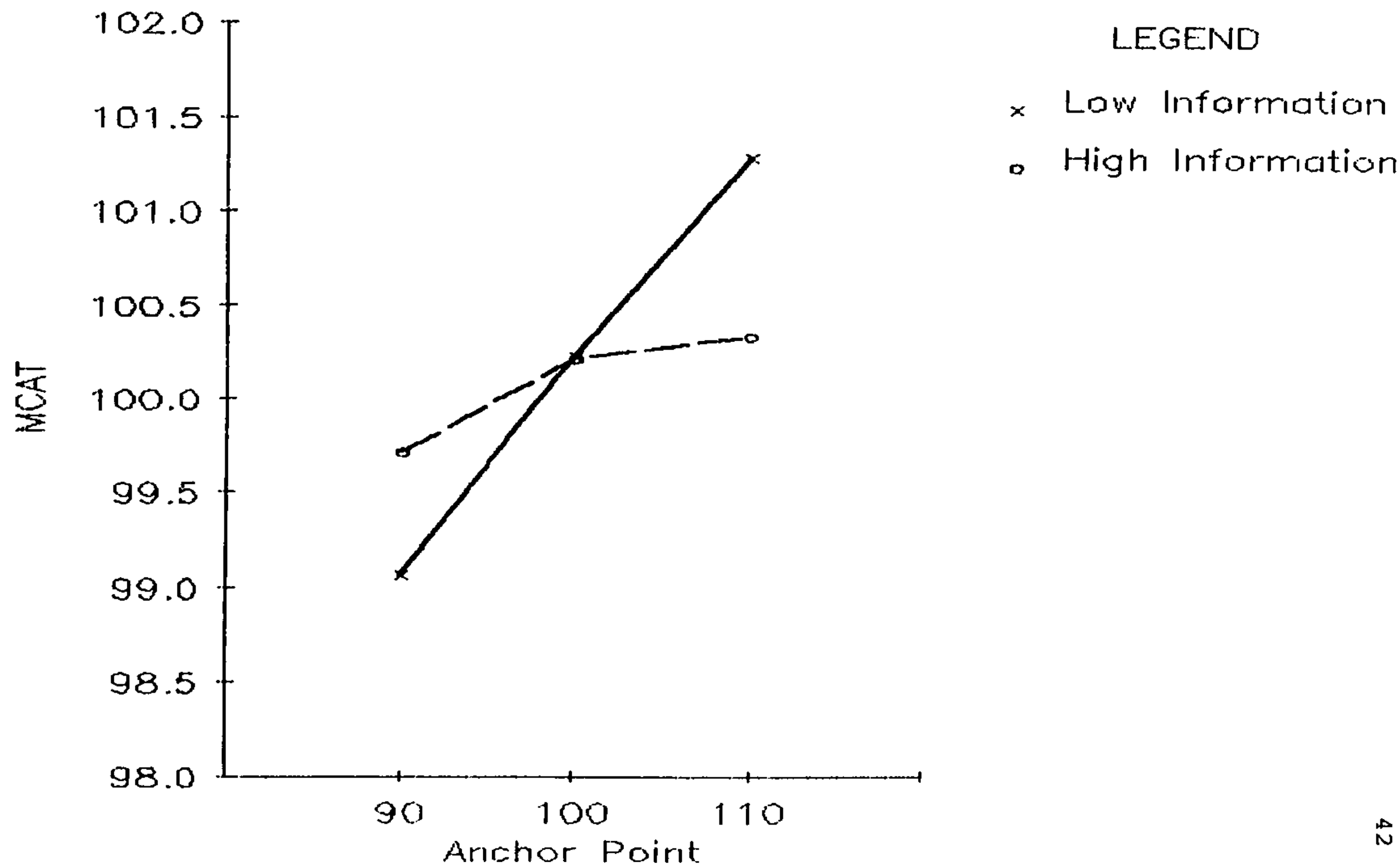
Order effects. Although the experimental design was carefully counterbalanced, subsidiary analyses were performed to determine if order of problem presentation confounded the experimental results. A five-way ANOVA (not reported) for the four dependent variables (MCAT, SCAT, CONJ and PTE) by variability, amount of information, anchor point, problem type and order showed that it mattered little whether the rainfall or the basketball problem was presented first or second. Judgments for the first problem were slightly more variable than for the second problem regardless of whether the low or high variability condition was presented first. Order did not systematically interact with any other factors.

Variability. First, variability affected the standard deviation of the categorical probability estimates ($F(1,1470) = 50.49$, $p < .01$). Problems in the low variability condition had a mean SCAT value of 8.19 (see Table 5.1), while problems in the high variability condition had a mean SCAT value of 9.57 (see Table 5.2).

Second, the results obtained indicate that variability affected subjects' confidence judgments ($F(1,1470) = 7.84$, $p < .01$). Problems in the low variability condition had a mean CONJ value of 3.56 (see Table 5.1) while problems in the high variability condition had a mean CONJ value of 3.37 (see Table 5.2). Therefore, subjects were more confident about their judgments in the low variability condition than in the high variability condition.

Anchor point by amount of information interaction. There was a significant ANCHOR by INFO interaction on MCAT values ($F(2,1470) = 5.39$, $p < .01$) that probably accounted for the overall effects of ANCHOR on MCAT. MCAT values as a function of INFO and ANCHOR are plotted in Figure 1. The results presented in Figure 1 and Table 8 show that anchor point affected MCAT values in the low

Figure 1. Mean Categorical Probability Estimates (MCAT) as a function of Amount of Information (INFO) and Anchor Point (ANCHOR)



information condition but not in the high information condition. The four-way ANOVA was broken into two separate three-way ANOVAs (TYPE by VAR by ANCHOR) by the levels of information (INFO) to assess the nature of the interaction. When subjects received small amounts of information, anchor point affected MCAT values ($F(2,745) = 19.99$, $p < .01$). The mean MCAT value for an anchor point of 90 was 99.06; the mean MCAT value for an anchor point of 100 was 100.22; and the mean MCAT value for an anchor point of 110 was 101.28. Tukey's HSD test revealed that the three means were all significantly different from one another ($p < .01$).

When subjects received large amounts of information, anchor point did not affect MCAT values ($F(2,747) = 1.86$, NS). The mean MCAT value for an anchor point of 90 was 99.71; the mean MCAT value for an anchor point of 100 was 100.21; an the mean MCAT value for an anchor point of 110 was 100.33. No significant differences among the three means were found. In the low information condition, mean MCAT values were pulled below the mean by an anchor point of 90; remained close to the mean for an anchor point of 100; and were pulled above the mean by an anchor point of 110. In the high information condition, however, mean MCAT values were close to the mean (100) for all three anchor points.

Anchor point. Although anchor point affected the mean categorical probability estimate overall ($F(2,1470) = 16.97$, $p < .01$), this result was probably due to the ANCHOR by INFO interaction.

Second, anchor point affected point estimates of the mean ($F(2,1470) = 20.90$, $p < .01$). Problems with an anchor point of 90 had a mean PTE value of 96.74; problems with an anchor point of 100 had a mean PTE value of 98.62; and problems with an anchor point of 110 had a mean PTE value of 99.23 (see Table 5.3). Post hoc analysis of the means using Tukey's HSD test revealed that an anchor point of 90 was significantly different from an anchor point of 100 or 110 ($p < .01$) but that anchor points of 100 and 110 were not significantly different from one another. Thus, an anchor point of 90 pulled PTE values well below the mean while PTE values for an anchor point of 100 remained below the mean and PTE values for an anchor point of 110 hovered slightly below the mean.

Variability by amount of information interaction. Although the results presented in Table 8 show that amount of information and variability interacted to affect SCAT values ($F(1,1470) = 9.89$, $p < .01$), this interaction was probably an artifact of problem construction. Individual problems were constructed so that they had a standard deviation of 1.20 (low information condition) and 1.25 (high information condition) in the low variability condition and standard deviations of 6.00 (low information condition) and 8.65 (high information condition) in the high variability condition. There was a difference in judged variability in the high variability condition but not in the low variability condition. The results obtained suggest that VAR accounted for the interaction because it affected SCAT but INFO did not (see Table 8).

Secondary Analyses

Gender affected the dependent measures, but age did not. Therefore, gender was the only subject characteristic to be further analyzed.

Gender. Gender affected some dependent variables by itself and in combination with problem type (i.e., problem type), but did not interact with any other independent variable. Table 9 provides the TYPE by GENDER ANOVA summary table for MCAT, SCAT, CONJ and PTE. Gender affected SCAT

Table 9

ANOVA Summary Table for the 4 Dependent Variables (MCAT, SCAT, CONJ, and PTE) by Problem Type (TYPE) and Gender (GENDER)

		MCAT		SCAT		CONJ		PTE	
Source:	df	SS	F	SS	F	SS	F	SS	F
TYPE	1	28.84	1.88	.17	.01	4.56	3.17	.22	.01
GENDER	1	14.78	.97	357.40	24.73*	307.23	213.50*	627.10	15.40*
TYPE X GENDER	1	37.19	2.43	17.42	1.21	16.61	11.54*	54.20	1.33
ERROR	1490	22798.00		21533.00		2144.10		60811.00	
TOTAL	1493	22865.00		21910.00		2469.70		61496.00	
R^2		.003		.017		.131		.040	

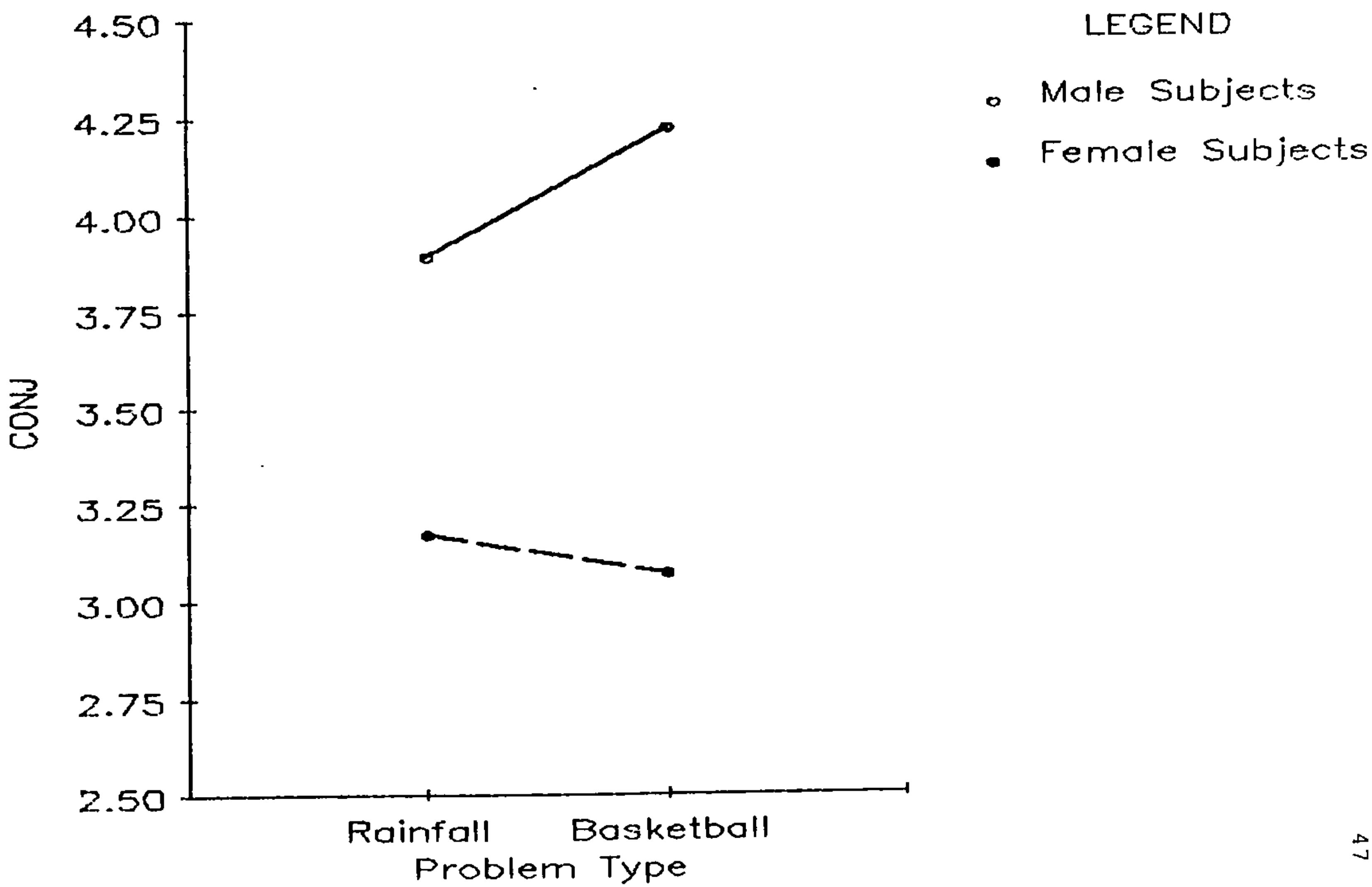
*$p < .01$

(F(1,1490) = 24.73, p < .0l), CONJ (F(1,1490) = 213.49, p < .01), and PTE (F(1,1490) = 15.38, p < .01). In addition, TYPE by GENDER interacted to affect CONJ (F(l,l490) = 11.54, p < .01). A five-way ANOVA for the four dependent variables by variability, amount of information, anchor point, problem type and gender was also performed. It is not reported here because it produced similar results. CONJ values as a function of TYPE and GENDER are plotted in Figure 2. Although there was a TYPE by GENDER interaction on CONJ, the effects of gender on CONJ are significant because male means are greater than female means for both problem types (see Figure 2).

Comparisons between the means (see Table 6) showed that male subjects had a mean SCAT value of 8.24, and female subjects had a mean SCAT value of 9.25. In addition, male subjects had a mean CONJ value of 4.06, while female subjects had a mean CONJ value of 3.11. Finally, male subjects had a mean PTE value of 99.05, while female subjects had a mean PTE value of 97.71. Mean differences for the three comparisons listed above and female means were statistically significant at p<.01 according to Tukey's HSD test.

Besides the main effect of gender on CONJ, TYPE and GENDER also combined to affect confidence judgments. The two-way ANOVA was broken into separate one way ANOVAs by TYPE across GENDER to more fully explore the interaction effect. TYPE affected CONJ values for male subjects (F(1,550) = 10.72, p < .01). The mean CONJ value for the rainfall problem was3.89 and the mean CONJ value for the basketball problem was 4.22. In contrast, TYPE did not affect CONJ for female subjects (F(1,940) = 1.76, NS). The mean CONJ value for the rainfall problem was 3.17, and the mean CONJ value for the basketball problem

Figure 2. Confidence Judgments (CONJ) as a function of Problem Type (TYPE) and Gender (GENDER)



was 3.07. The results presented above indicate that males were more confident about their judgments of the basketball problem than the rainfall problem, while females were equally confident about their judgments for both problems.

In summary, gender affected several dependent variables in this study. Males and females had differences in standard deviations for their subjective probability distributions, confidence judgments and point estimates of the mean from memory. Gender and problem type combined to affect confidence judgments, with males being more confident about the sports problem than the weather problem and females being equally confident about both problems. It appears that gender systematically affects subjective probability estimates, point estimates, and confidence judgments at least in this study.

Summary of Results

Table 10 summarizes the main findings for the study for all four dependent variables (MCAT, SCAT, CONJ, & PTE). MCAT, the mean categorical probability estimate, was primarily affected by the INFO by ANCHOR interaction. Although ANCHOR affected MCAT overall, closer examination revealed that this result was due to the INFO by ANCHOR interaction. Anchor point affected MCAT in the low information condition but not in the high information condition.

SCAT, the standard deviation of the categorical probability estimates, was affected by VAR and by GENDER. Subjects had larger standard deviations for their subjective probability distributions in the high variability condition than in the low variability condition. Females had larger standard deviations for their subjective probability distributions than did males.

Table 10

Summary of Main Findings for all 4 Dependent Variables

(MCAT,SCAT, CONJ, and PTE)

	MCAT	SCAT	CONJ	PTE
Source:				
TYPE				
VAR		*	*	
INFO				
ANCHOR	*			*
GENDER		*	*	*
INFO X ANCHOR	*			
TYPE X GENDER			*	

* indicates that F-Value was significant at $p < .01$

CONJ, the subjects' confidence in their subjective probability judgments, was affected by VAR, GENDER, and the TYPE by GENDER interaction. Subjects had greater confidence in their judgments for the low variability condition than for the high variability condition. Males were more confident than females for both the rainfall and basketball problems. Problem type and gender interacted to affect CONJ. Males were more confident about the basketball problem than about the rainfall problem, while females were equally confident about both problems.

PTE, the point estimate of the mean from memory, was affected by ANCHOR and by GENDER. An anchor point of 90 pulled PTE values well below the actual mean (99.8), but PTE values remained slightly below the mean for both an anchor point of 100 and an anchor point of 110. Males made point estimates that were closer to the actual mean given in the problem than did females.

CHAPTER FOUR

Discussion

The discussion is divided into three sections:

1. The relationship of results to hypotheses;
2. Implications for research on subjective probability judgments.
3. Suggestions for future research studies.

Relationship of results to hypotheses

The purpose of the present study was to see how people would respond to a decision problem that manipulated several characteristics of a data distribution. A number of issues in subjective probability estimation were examined, the most important of which is the normative appropriateness of subjects' judgments. The discussion of results is organized by independent variables and/or their interactions.

Variability. Hypothesis 1 dealt with effects of VAR on SCAT. It stated that subjective probability distributions for problems consisting of data with high variability will have larger judged standard deviations than problems consisting of data with low variability. Some problems had a small standard deviation (low variability condition), while other problems had a large standard deviation (high variability condition). It was proposed that the subjective probability distributions for problems with large standard deviations would be more variable than subjective probability distributions for problems with small standard deviations. The results presented in Table 8 indicated that problem variability did affect the standard deviation of the subjective probability estimates (SCAT). Problems with small standard deviations had a mean SCAT value of 8.19, while problems with large standard deviations had a mean SCAT value of 9.57 (see Table 5.3). These results showed that people can detect relative differences in problem variability and that they can make use of these differences when making probability judgments.

It appears that subjects made judgments that were normatively appropriate. Distributions with large standard deviations should indeed be judged as being more variable than distributions with small standard deviations. However, subjects only detected relative differences in variability and not absolute differences. They accurately estimated standard deviations in the high variability condition but overestimated standard deviations in the low variability condition. It would seem that subjects perceived a standard deviation of 1.50 as small and a standard deviation of 10 as moderate but not large for numbers averaging about 100. The overestimation of standard deviations in the low variability condition may not be entirely due to subject error. Actual SCAT sizes may have been dependent on the category boundaries used for the subjective probability judgments. Consequently, the overestimation bias for the low variability condition may have resulted from the selection of arbitrary category boundaries and not from subject estimation errors.

Hypothesis 2 dealt with the effects of VAR on CONJ. It stated that problems with large standard deviations will yield lower confidence judgments than problems with small standard deviations. Subjects were more confident about problems in the low variability condition (mean CONJ value of 3.56) than about problems in the high variability condition (mean CONJ value of 3.37). The results obtained imply that a network of relationships exists between variability, uncertainty and confidence judgments. Variability and uncertainty are closely related but not synonymous concepts (Tversky and Kahneman, 1982). Variability can be defined as the dispersion of data in a distribution. Uncertainty is present in many types of decisions and can be defined as the inability to assign specific probabilities to outcomes. Different degrees of uncertainty can be distinguished. For some tasks, individuals may not know exact probabilities of outcomes but may have enough information to make estimates for their range such as .10 to .30. For other tasks, it may not be possible to assign any probability estimates for outcomes. Such tasks would be characterized as having greater uncertainty.

Kahneman and Tversky (1982) found that uncertainty judgments were based on the distributional characteristics of data, and Yates and Zukowski (1976) reported that uncertainty increased as the variability of the subjective probability distribution increased. Taken together, these findings indicate that high variability is associated with high uncertainty. Slovic, Fischoff, and Lichtenstein (1980) found that high levels of uncertainty are associated with lower confidence judgments. Confidence judgments represent an individual's uncertainty about a prediction, estimate or inference that has already been made (Kahneman and Tversky, 1982). They are always made in reference to the individual's confidence in the accuracy of his/her estimates.

The results obtained revealed that female subjects had higher variability and less confidence than male subjects and it can be inferred that they also had higher uncertainty. Moreover, subjects were less confident about problems in the high variability condition than about problems in the low variability condition. On the basis of previous research and on the findings for this study, high variability is associated with high uncertainty, and high uncertainty is associated with low confidence. Using uncertainty as a mediating variable, transitivity rules suggest that high variability is associated with low confidence. More attention needs to be given to the interrelationships between variability, uncertainty and confidence because of the role that they play in judgment and decision processes.

Amount of information. Hypothesis 3 dealt with the effects of INFO on CONJ. It stated that problems based on 25 years of data will yield higher confidence judgments than problems based on five years of data. Although amount of information interacted with anchor point to affect MCAT values, as has been noted, it did not affect any of the dependent measures by itself. One possible explanation for the result obtained is that other task variables may have had a greater effect on people's judgments. Nevertheless, the role of amount of information in decision making should not be slighted. Previous research has revealed that the amount of task information, number of alternatives, number of outcomes and the way information is presented all affect individual choice (Nisbett and Ross, 1980; Payne, 1982). Perhaps, amount of information would have had a greater impact on judgments if other types of tasks or dependent measures had been used.

Anchor point by amount of information interaction. Hypothesis 4 dealt with the effects of the interaction of ANCHOR and INFO on MCAT. It stated that mean estimates will be biased in the direction of the anchor point for the low information condition but not the high information condition. This hypothesis was confirmed. The results presented in Figure 1 and Table 8 indicate that anchor point affected MCAT values in the low information condition but not in the high information condition. In the low information condition, mean MCAT values were pulled below the problem mean by an anchor point of 90, remained near the mean for an anchor point of 100 and were pulled above the mean for an anchor point of 110. In contrast anchor point did not affect MCAT values in the high information condition and mean MCAT values were much more tightly clustered than the means in the low information condition.

One explanation for the results obtained is that the effects of highly discrepant values are greatest with small data sets. Just as an arithmetic mean can be highly influenced by extreme values, subjective probability judgments are also susceptible to the effects of discrepant data points (Kahneman and Tversky, 1979). In the present study, anchor points of 90 and 110 were discrepant from the problem mean of close to 100. When only five other data points were available, these anchors strongly affected MCAT values and pulled them in the direction of the anchor point. It appears as if subjects used the anchor points as an additional data point and based their mean estimates on six rather than five pieces of information.

However, the results obtained were different when 25 years of data were available. Anchor point had little or no effect on mean MCAT values in this condition. It again appears that the subjects used the anchor points as an additional data point in the high information condition. In so doing, they based their mean estimates on 26 rather than 25 data points. Discrepant anchor points seem to have their greatest effects when samples are small. The effect of more data is to reduce the impact of extreme values. This result may indicate that subjects made normatively appropriate judgments.

It is worthwhile noting that there was no anchor point by amount of information interaction for PTE values. The amount of data had no direct effect on judgments made from memory and subjects did not make judgments normatively appropriate to the data given. One explanation for this finding is that subjects did not take into account the fact that larger amounts of data reduce the impact of extreme values. As a result, discrepant anchor points, particularly those below the mean, had a greater affect on PTE values. Because PTE values were based on judgments made from memory, it is possible that subjects forgot whether their mean estimates were based on small or large amounts of data.

Gender by problem type interaction. Hypothesis 5 dealt with the interaction of GENDER and TYPE on CONJ. It stated that men will be more confident about thier judgments for the basketball problem than the rainfall problem while women will be equally confident about their judgments for both problems. This hypothesis was confirmed. Male subjects had a mean confidence judgment (CONJ) of 3.89 for the rainfall problem and a mean CONJ value of 4.22 for the basketball problem. In contrast, female subjects had a mean CONJ value of 3.17 for the rainfall problem and a mean CONJ value of 3.07 for the basketball problem. Therefore, males were more confident about the basketball problem than the rainfall problem while females were equally confident about both problems.

These findings may be due to gender differences in self confidence and to differences in task content. In general, men appear to have more self confidence than women (Hyde and Rosenberg, 1980). However, Lenney (1977) found that some limitations need to be placed on this general result. Her research showed that women's confidence judgments are dependent upon the gender appropriateness of the task. If women are given a gender appropriate task, then their confidence judgments are similar to men.

In the context of the present study, it seems reasonable to assume that the basketball problem represents a male specific task and the rainfall problem represents a gender neutral task. Consequently, there was no female specific task. It could be argued that the results obtained were a function of task content. Research has suggested that people often base their judgments on their knowledge structures, schemas and scripts (Kelley, 1972; Abelson, 1978). Because most males have participated in or viewed a basketball game and many females have not, males felt more knowledgeable and more confident about the basketball problem than about the rainfall problem. In contrast, females felt equally knowledgeable and equally confident about both problems. Perhaps different results would be obtained if a female specific task would have been used in the study.

Implications for research on subjective probability judgments.

The purpose of the study was to see how various characteristics of the data distribution affected subjective probability judgments, confidence judgments and point estimates. Unlike previous research on decision making (cf. Tversky and Kahneman, 1971), the present research simultaneously manipulated several characteristics of the data distribution; thereby, allowing both two and three way interactions to be examined. Moreover, the present study used multiple dependent measures as opposed to the single dependent measures (usually a mean estimate) used in previous research.

In examining the characteristics of subjects' subjective probability judgments, there were two main issues. First, how normative were subjects' judgments? Although other studies have addressed this issue (Winkler, 1971; Peterson and Beach, 1967), the present study examined whether subjects could make normative judgments when several characteristics of the data distribution were simultaneously manipulated. A review of the decision making literature revealed some studies which reported using multiple independent and dependent variables (Pratt, 1982), but they were in different research contexts than the present study. Second, how confident did subjects feel about their probability judgments? Although other researchers have studied confidence judgments in the context of probability estimation (Lichtenstein, Fischoff and Phillips, 1982), the present study attempted to define the relationship between variability and confidence judgments in a more complex situation.

Examination of the results revealed several main effects and interaction effects. However, a greater number of significant findings was expected, especially given the large number of independent variables. The results indicated that most individual independent variables affected only one dependent measure. Moreover, each main effect for an independent variable was associated with a different dependent measure. Each significant two-way interaction also involved only one dependent measure.

Because different independent variables affected different dependent measures, it appears that different distributional characteristics function independently of one another. One implication of this finding is that the effects for a particular distributional characteristic must be interpreted individually in a decision task.

The results obtained for the present study suggest that subjects recognized, at least implicitly, that statistical principles affect their judgments and that distributional attributes of data are important in making judgments. For example, the variability of the data distributions affected the subjects' probability judgments and subjects seemed to recognize that extreme values affect mean responses less if the sample size is large. Therefore, the results for this study indicate that people can make accurate (i.e., normatively appropriate) as well as erroneous judgments.

In addition, the results linking variability and confidence, and uncertainty and confidence suggest that variability, uncertainty and confidence are theoretically connected. Variability and uncertainty covary directly (Kahneman and Tversky, 1982), as do uncertainty and confidence (Lichtenstein, Fischoff and Phillips, 1982). This leads to a possible linkage of the three constructs as follows:

Variability ---> (Uncertainty) ---> Confidence Judgments

That is, the observed relationship between variability and confidence operates through task uncertainty as a mediating variable. Further study should indicate whether this finding is generalizable to other research contexts.

<u>Suggestions for future research studies</u>

Based on the findings obtained in this study, two studies suggest themselves. They are presented below.

<u>Gender specificity of task</u>. Although male subjects received a male specific task in this study, female subjects did not receive a female specific task. A future study might have three tasks: male specific, female specific and gender neutral. It is hypothesized that males would be more confident than females on a male specific task and a gender neutral task, but that they would be equally or less confident than females about a female specific task.

Exploration of the relationship between variability, uncertainty and confidence judgments. The results of the present study implied that a network of relationships existed between variability, uncertainty and confidence. A further study might examine how variability and uncertainty together affect confidence judgments. With a scenario similar to the rainfall problem, both variability and uncertainty of the problem could be manipulated. It is hypothesized that high levels of both variability and uncertainty would yield very low confidence judgments, while low levels of both variability and uncertainty would yield very high confidence judgments.

Summary

The major findings of the study were:

1. The mean of the subjective probability distribution was affected by the interaction of amount of information and anchor point.

2. The standard deviation of the subjective probability distribution was affected by problem variability and by gender.

3. Confidence judgments were affected by problem variability, gender and a problem type by gender interaction.

4. On the whole, subjects tended to make normatively appropriate probability judgments.

References

Alpert, M., & Raiffa, H. (1982). A progress report on the training of probability assessors. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases (pp. 294-305). Cambridge: Cambridge University Press.

Anastasi, A. (1958). Differential psychology (3rd ed.) NewYork: Macmillan.

Bar-Hillel, M. (1979). The role of sample size in sample evaluation. Organizational Behavior and Human Performance, 24, 245-257.

Christensen-Szalanski, J. J., & Beach, L. R. (1984). The citation bias: Fad and fashion in the judgment and decision literature. American Psychologist, 39, 75-78.

Cantor, N., Pittman, T., & Jones, E. (1982). Choice and attitude attributions: The influence of constraint information on attributions across levels of generality. Social Cognition, 1, 1-20.

Cliff, N. (1959). Adverbs as multipliers. Psychological Review, 66, 27-44.

Cohen, J., Dearnaley, E. J., & Hansel, C. E. M. (1956). The addition of subjective probabilities: The summation of estimates of success and failure. Acta Psychologica, 12, 37l-380.

Cohen, J., Deamaley, E. J., & Hansel, C. E. M. (1957). Measures of subjective probability: Estimates of success in performance in relation to size of task. British Journal of Psychology, 48, 272-275.

Cohen, J., & Hansel, C. E. M. (1955). The idea of a distribution. British Journal of Psychology, 46, 111-121.

Cohen, J., & Hansel, C. E. M. (1955). The idea of a distribution. British Journal of Psychology, 46, 111-121.

Coombs, C. H., Dawes, R. M., & Tversky, A. (1970). Mathematical psychology: An elementary introduction. Englewood Cliffs, NJ: Prentice-Hall.

Devore, J. (1987). Probability and statistics for engineering and the sciences. (2nd ed.). Monterey, CA: Brooks/Cole.

Edwards, W. (1954). The theory of decision making. Psychological Bulletin, 51, 380-417.

Edwards, W. (196l). Behavioral decision theory. Annual Review of Psychology, 12, 473-498.

Einhorn, H. J. (197l). Use of nonlinear, noncompensatory models as a function of task and amount of information. Organizational Behavior and Human Performance, 6, 127-139.

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence in the illusion of validity. Psychological Review, 85, 395-416.

Einhorn, H. J., & Hogarth, R. M. (198l). Behavioral decision theory: Processes theory: Processes of judgment and choice. Annual Review of Psychology, 32, 53-88.

Funder, D. (1987). Errors and mistakes: Evaluating the accuracy of social judgment, Psychological Bulletin, 101, 75-90.

Hogarth, R. M. (198l). Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics. Psychological Bulletin, 90, 197-217.

Holloway, C. (1979). Decision making under uncertainty: Models and choices. Englewood Cliffs, NJ: Prentice-Hall.

Howell, W. & Burnett, S. (1978) . Uncertainty measurement : a cognitive taxonomy. Organizational Behavior and Human Performance, 22, 45-68.

Hyde, J., & Rosenberg, B. G. (1980). Half the human experience: The psychology of women (2nd ed.). Lexington, MA: D.C. Heath.

Kahneman, D., Slovic, P., & Tversky, A. (1982). Judgment under uncertainty: Heuristics and biases. Cambridge:Cambridge University Press.

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

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. Econometrica, 47, 263-291.

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. Cognition, 11, 123-141.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. American Psychologist, 39, 341-350.

Kelley, H. H., (1972). Causal schemata and the attribution process. In E. E. Jones (Ed.), Attribution: Perceiving the causes of behavior. Morristown, NJ: General Learning Press.

Kelley, H. H., (1973). The process of causal attribution. American Psychologist, 28, 107-128.

Keppel, G. (1982). Design and analysis: A researcher's handbook (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Kleiman, S. (1983). The effects of probability level and sample size on subjective probability distributions. Unpublished master's thesis, University of Windsor,Windsor, ONT.

Lehman, D., Lempert, R., & Nisbett, R. (1988). The effects of graduate training on reasoning: Formal discipline and thinking about everday-life events. American Psychologist, 43, 431-442.

Lenney, E. (1977). Women's self-confidence in achievement settings. Psychological Bulletin, 84, 1-13.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases. Cambridge: Cambridge University Press.

MacCoby, E. E., & Jacklin, C. N. (1974). The psychology of sex differences. Stanford, CA: Stanford University Press.

Murphy, A., & Winkler, R. (1974). Probability forecasts: A survey of National Weather Service forecasters. Bulletin of the American Meteorological Society, 55, 1449-1453.

Neale, M, Huber, V., & Northcraft, G. (1987). The framing of negotiations: Contextual versus task frames. Organizational Behavior and Human Decision Processes, 39, 228-241.

Newell, A., & Simon, H. A. (1972). Human problem solving. Englewood Cliffs, NJ: Prentice-Hall.

Nisbett, R., Fong, G., Lehman, D., & Cheng, P. (1987). Teaching reasoning. Science, 238, 625-631.

Nisbett, R. E., & Ross, L. (1980). Human inference: Strategies and shortcomings of social judgment. Englewood Cliffs, NJ: Prentice-Hall.

Payne, J. W. (1982). Contingent decision behavior. Psychological Bulletin 92, 382-402.

Peterson, C. R., & Beach, L. R. (1967). Man as intuitive statistician. Psychological Bulletin, 68, 29-46.

Pitz, G. P., & Sachs, N. J. (1984). Judgment and decision:Theory and application. Annual Review of Psychology, 35,139-163.

Pratt, J. (1982). Post-Cognitive structure: Its determinants and relationship to perceived information use and predictive accuracy. Journal of Accounting Research, 20, 189-209.

Seaver, D., von Winterfeldt, D., & Edwards, W. (1978).Eliciting subjective probability distributions on continuous variables. Organizational Behavior and HumanPerformance, 21, 379-391.

Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior andHuman Performance, 6, 649-744.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1976). Cognitive processes and societal risk taking. In J. S. Carroll & J. W. Payne (Eds.), Cognition and social behavior. Hillsdale, NJ: Erlbaum.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. Annual Review of Psychology, 28, 139-163.

Solso, R. (1979). Cognitive psychology. New York: Harcourt Brace Jovanovich.

Starr, M. W., & Kleiman, S. (1985). The under/over bettor bias: An artifact of subjective probability estimation. In W. R. Eadington (Ed.) The Gambling Studies: Proceedings of the 6th National Conference on Gambling and Risk Taking. (pp. 13-41). Reno, NV: University of Nevada, Reno.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. Psychological Bulletin, 76, 105-110.

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

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

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. Science, 211, 453-458.

Tversky, A., & Kahneman, D. (1982). Evidential impact of base rates. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics andbiases. Cambridge: Cambridge University Press.

Tversky, A., & Kahneman, D. (1982). Variants of uncertainty. Cognition, 11, 143-157.

Wallsten, T. (1980) (Ed.). Cognitive processes in choice and decision behavior. Hillsdale NJ: Erlbaum.

Winkler, R. (1971). Probabilistic prediction: Some experimental results. Journal of the AmericanStatistical Association, 66, 675-685.

Wright, G. (1985). (Ed.). Behavioral decision making. New York: Plenum Press.

Yates, J. F., & Zukowski, L. G. (1976). Characterization of ambiguity in decision making. Behavioral Science, 21,19-25.

Appendix A

All Versions of the Experimental Questionnaire

DECISION PROCESSES STUDY

(Do NOT write your name on this questionnaire)

Sex: ______________________________ Age, in years: _____________

Town/City of residence: _____________________

Number of years of university: ______________

Decisions about various events must be made daily. For example, individuals are concerned about the likelihood that it will rain today or the likelihood that the Detroit Tigers will win today's baseball game. I believe that it is interesting to study how individuals make decisions. On the following pages, you are asked to make predictions about meteorological events and about sports events. Please read the problems carefully and then indicate your response on the blanks provided. Answer the problems to the best of your ability. Thank you for your cooperation.

ANNUAL RAINFALL FORECAST

VAR = LOW
INFO = LOW
ANCHOR = 90

Meteorologists in Buenos Aires, Argentina decided to examine the pattern of rainfall in Buenos Aires. The Archives of the Argentinian Meteorological Society showed that the measurable annual rainfall (in cm.) in Buenos Aires, Argentina for the years 1980 to 1984 was:

1980 : 98.4

1981 : 99.1

1982 : 101.5

1983 : 100.4

1984 : 99.7

In 1985, the annual rainfall in Buenos Aires was 90 cm.

Please continue to next page.

ANNUAL RAINFALL FORECAST

VAR = LOW
INFO = LOW
ANCHOR = 100

Meteorologists in Buenos Aires, Argentina decided to examine the pattern of rainfall in Buenos Aires. The _Archives of the Argentinian Meteorological Society_ showed that the measurable annual rainfall (in cm.) in Buenos Aires, Argentina for the years 1980 to 1984 was:

1980 : 98.4

1981 : 99.1

1982 : 101.5

1983 : 100.4

1984 : 99.7

In 1985, the annual rainfall in Buenos Aires was 100 cm.

Please continue to next page.

ANNUAL RAINFALL FORECAST

VAR = LOW
INFO = LOW
ANCHOR = 110

Meteorologists in Buenos Aires, Argentina decided to examine the pattern of rainfall in Buenos Aires. The Archives of the Argentinian Meteorological Society showed that the measurable annual rainfall (in cm.) in Buenos Aires, Argentina for the years 1980 to 1984 was:

1980 : 98.4

1981 : 99.1

1982 : 101.5

1983 : 100.4

1984 : 99.7

In 1985, the annual rainfall in Buenos Aires was 110 cm.

Please continue to next page.

ANNUAL RAINFALL FORECAST

VAR = HIGH
INFO = LOW
ANCHOR = 90

Meteorologists in Buenos Aires, Argentina decided to examine the pattern of rainfall in Buenos Aires. The Archives of the Argentinian Meteorological Society showed that the measurable annual rainfall (in cm.) in Buenos Aires, Argentina for the years 1980 to 1984 was:

1980 : 95.6

1981 : 93.2

1982 : 103.4

1983 : 107.1

1984 : 99.7

In 1985, the annual rainfall in Buenos Aires was 90 cm.

Please continue to next page.

ANNUAL RAINFALL FORECAST

VAR = HIGH
INFO = LOW
ANCHOR = 100

Meteorologists in Buenos Aires, Argentina decided to examine the pattern of rainfall in Buenos Aires. The Archives of the Argentinian Meteorological Society showed that the measurable annual rainfall (in cm.) in Buenos Aires, Argentina for the years 1980 to 1984 was:

1980 : 95.6

1981 : 93.2

1982 : 103.4

1983 : 107.1

1984 : 99.7

In 1985, the annual rainfall in Buenos Aires was 100 cm.

Please continue to next page.

ANNUAL RAINFALL FORECAST

VAR = HIGH
INFO = LOW
ANCHOR = 110

Meteorologists in Buenos Aires, Argentina decided to examine the pattern of rainfall in Buenos Aires. The Archives of the Argentinian Meteorological Society showed that the measurable annual rainfall (in cm.) in Buenos Aires, Argentina for the years 1980 to 1984 was:

1980 : 95.6

1981 : 93.2

1982 : 103.4

1983 : 107.1

1984 : 99.7

In 1985, the annual rainfall in Buenos Aires was 110 cm.

Please continue to next page.

ANNUAL RAINFALL FORECAST

VAR = LOW
INFO = HIGH
ANCHOR = 90

Meteorologists in Buenos Aires, Argentina decided to examine the pattern of rainfall in Buenos Aires. The Archives of the Argentinian Meteorological Society showed that the measurable annual rainfall (in cm.) in Buenos Aires, Argentina for the years 1960 to 1984 was:

1960 : 102.0	1969 : 100.5	1977 : 98.0
1961 : 101.1	1970 : 98.4	1978 : 98.1
1962 : 98.1	1971 : 99.4	1979 : 100.8
1963 : 100.3	1972 : 99.8	1980 : 101.0
1964 : 101.6	1973 : 98.3	1981 : 98.6
1965 : 98.2	1974 : 101.7	1982 : 100.2
1966 : 99.2	1975 : 100.2	1983 : 100.4
1967 : 101.2	1976 : 99.6	1984 : 99.7
1968: 99.0		

In 1985, the annual rainfall in Buenos Aires was 90 cm.

Please continue to the next page.

ANNUAL RAINFALL FORECAST

VAR = LOW
INFO = HIGH
ANCHOR = 100

Meteorologists in Buenos Aires, Argentina decided to examine the pattern of rainfall in Buenos Aires. The Archives of the Argentinian Meteorological Society showed that the measurable annual rainfall (in cm.) in Buenos Aires, Argentina for the years 1960 to 1984 was:

1960 : 102.0	1969 : 100.5	1977 : 98.0
1961 : 101.1	1970 : 98.4	1978 : 98.1
1962 : 98.1	1971 : 99.4	1979 : 100.8
1963 : 100.3	1972 : 99.8	1980 : 101.0
1964 : 101.6	1973 : 98.3	1981 : 98.6
1965 : 98.2	1974 : 101.7	1982 : 100.2
1966 : 99.2	1975 : 100.2	1983 : 100.4
1967 : 101.2	1976 : 99.6	1984 : 99.7
1968: 99.0		

In 1985, the annual rainfall in Buenos Aires was 100 cm.

Please continue to the next page.

ANNUAL RAINFALL FORECAST

VAR = LOW
INFO = HIGH
ANCHOR = 110

Meteorologists in Buenos Aires, Argentina decided to examine the pattern of rainfall in Buenos Aires. The Archives of the Argentinian Meteorological Society showed that the measurable annual rainfall (in cm.) in Buenos Aires, Argentina for the years 1960 to 1984 was:

1960 : 102.0	1969 : 100.5	1977 : 98.0
1961 : 101.1	1970 : 98.4	1978 : 98.1
1962 : 98.1	1971 : 99.4	1979 : 100.8
1963 : 100.3	1972 : 99.8	1980 : 101.0
1964 : 101.6	1973 : 98.3	1981 : 98.6
1965 : 98.2	1974 : 101.7	1982 : 100.2
1966 : 99.2	1975 : 100.2	1983 : 100.4
1967 : 101.2	1976 : 99.6	1984 : 99.7
1968: 99.0		

In 1985, the annual rainfall in Buenos Aires was 110 cm.

Please continue to the next page.

ANNUAL RAINFALL FORECAST

VAR = HIGH
INFO = HIGH
ANCHOR = 90

Meteorologists in Buenos Aires, Argentina decided to examine the pattern of rainfall in Buenos Aires. The _Archives of the Argentinian Meteorological Society_ showed that the measurable annual rainfall (in cm.) in Buenos Aires, Argentina for the years 1960 to 1984 was:

1960 : 110.0	1969 : 102.9	1977 : 99.1
1961 : 112.4	1970 : 85.0	1978 : 94.8
1962 : 92.9	1971 : 96.7	1979 : 108.8
1963 : 107.6	1972 : 90.4	1980 : 103.1
1964 : 114.8	1973 : 87.4	1981 : 96.2
1965 : 87.0	1974 : 100.3	1982 : 106.7
1966 : 93.4	1975 : 111.2	1983 : 108.1
1967 : 104.7	1976 : 86.7	1984 : 99.7
1968: 95.3		

In 1985, the annual rainfall in Buenos Aires was 90 cm.

Please continue to the next page.

ANNUAL RAINFALL FORECAST

VAR = HIGH
INFO = HIGH
ANCHOR = 100

Meteorologists in Buenos Aires, Argentina decided to examine the pattern of rainfall in Buenos Aires. The Archives of the Argentinian Meteorological Society showed that the measurable annual rainfall (in cm.) in Buenos Aires, Argentina for the years 1960 to 1984 was:

1960 : 110.0	1969 : 102.9	1977 : 99.1
1961 : 112.4	1970 : 85.0	1978 : 94.8
1962 : 92.9	1971 : 96.7	1979 : 108.8
1963 : 107.6	1972 : 90.4	1980 : 103.1
1964 : 114.8	1973 : 87.4	1981 : 96.2
1965 : 87.0	1974 : 100.3	1982 : 106.7
1966 : 93.4	1975 : 111.2	1983 : 108.1
1967 : 104.7	1976 : 86.7	1984 : 99.7
1968: 95.3		

In 1985, the annual rainfall in Buenos Aires was 100 cm.

Please continue to the next page.

ANNUAL RAINFALL FORECAST

VAR = HIGH
INFO = HIGH
ANCHOR = 110

Meteorologists in Buenos Aires, Argentina decided to examine the pattern of rainfall in Buenos Aires. The Archives of the Argentinian Meteorological Society showed that the measurable annual rainfall (in cm.) in Buenos Aires, Argentina for the years 1960 to 1984 was:

1960 : 110.0	1969 : 102.9	1977 : 99.1
1961 : 112.4	1970 : 85.0	1978 : 94.8
1962 : 92.9	1971 : 96.7	1979 : 108.8
1963 : 107.6	1972 : 90.4	1980 : 103.1
1964 : 114.8	1973 : 87.4	1981 : 96.2
1965 : 87.0	1974 : 100.3	1982 : 106.7
1966 : 93.4	1975 : 111.2	1983 : 108.1
1967 : 104.7	1976 : 86.7	1984 : 99.7
1968: 95.3		

In 1985, the annual rainfall in Buenos Aires was 110 cm.

Please continue to the next page.

ANNUAL RAINFALL FORECAST

Assume that you are given the task of predicting the annual rainfall (in cm.) in Buenos Aires for 1986. Based on the information given on the preceding page, please estimate the probability that each of the following categories will occur (in %).

	Probability Estimate (in %)
In 1986, less than 90 cm. of rain will fall in Buenos Aires, Argentina.	________
In 1986, 90 to 110 cm. of rain will fall in Buenos Aires, Argentina.	________
In 1986, greater than 110 cm. of rain will fall in Buenos Aires, Argentina.	________
	Total = 100%

Note: Your probability estimates should add up to about 100% because the categories include all possibilities.

Please place an X at a point on the scale below which best describes your confidence in the accuracy of your probability estimates above.

________	________	________	________	________	________
Not At All Confident	Slightly Confident	Rather Confident	Quite Confident	Very Confident	Extremely Confident

Please continue to the next page.

SPORTS SURVEY

VAR = LOW
INFO = LOW
ANCHOR = 90

A group of sportswriters in Boston, Massachusetts decided to examine the average points per game that the Boston Celtics scored over the past few seasons. Pro Basketball Digest showed that the average points per game scored by the Boston Celtics for the years 1980 to 1984 was:

1980 : 98.4

1981 : 99.1

1982 : 101.5

1983 : 100.4

1984 : 99.7

In 1985, the Boston Celtics scored an average of 90 points per game.

Please continue to the next page.

SPORTS SURVEY

VAR = LOW
INFO = LOW
ANCHOR = 100

A group of sportswriters in Boston, Massachusetts decided to examine the average points per game that the Boston Celtics scored over the past few seasons. Pro Basketball Digest showed that the average points per game scored by the Boston Celtics for the years 1980 to 1984 was:

1980 : 98.4

1981 : 99.1

1982 : 101.5

1983 : 100.4

1984 : 99.7

In 1985, the Boston Celtics scored an average of 100 points per game.

Please continue to the next page.

SPORTS SURVEY

VAR = LOW
INFO = LOW
ANCHOR = 110

A group of sportswriters in Boston, Massachusetts decided to examine the average points per game that the Boston Celtics scored over the past few seasons. Pro Basketball Digest showed that the average points per game scored by the Boston Celtics for the years 1980 to 1984 was:

1980 : 98.4

1981 : 99.1

1982 : 101.5

1983 : 100.4

1984 : 99.7

In 1985, the Boston Celtics scored an average of 110 points per game.

Please continue to the next page.

SPORTS SURVEY

VAR = HIGH
INFO = LOW
ANCHOR = 90

A group of sportswriters in Boston, Massachusetts decided to examine the average points per game that the Boston Celtics scored over the past few seasons. Pro Basketball Digest showed that the average points per game scored by the Boston Celtics for the years 1980 to 1984 was:

1980 : 95.6

1981 : 93.2

1982 : 103.4

1983 : 107.1

1984 : 99.7

In 1985, the Boston Celtics scored an average of 90 points per game.

Please continue to the next page.

SPORTS SURVEY

VAR = HIGH
INFO = LOW
ANCHOR = 100

A group of sportswriters in Boston, Massachusetts decided to examine the average points per game that the Boston Celtics scored over the past few seasons. Pro Basketball Digest showed that the average points per game scored by the Boston Celtics for the years 1980 to 1984 was:

1980 : 95.6

1981 : 93.2

1982 : 103.4

1983 : 107.1

1984 : 99.7

In 1985, the Boston Celtics scored an average of 100 points per game.

Please continue to the next page.

SPORTS SURVEY

VAR = HIGH
INFO = LOW
ANCHOR = 110

A group of sportswriters in Boston, Massachusetts decided to examine the average points per game that the Boston Celtics scored over the past few seasons. Pro Basketball Digest showed that the average points per game scored by the Boston Celtics for the years 1980 to 1984 was:

1980 : 95.6

1981 : 93.2

1982 : 103.4

1983 : 107.1

1984 : 99.7

In 1985, the Boston Celtics scored an average of 110 points per game.

Please continue to the next page.

SPORTS SURVEY

VAR = LOW
INFO = HIGH
ANCHOR = 90

A group of sportswriters in Boston, Massachusetts decided to examine the average points per game that the Boston Celtics scored over the past few seasons. Pro Basketball Digest showed that the average points per game scored by the Boston Celtics for the years 1960 to 1984 was:

1960 : 102.1	1969 : 100.5	1977 : 98.0
1961 : 101.1	1970 : 98.4	1978 : 98.1
1962 : 98.1	1971 : 99.4	1979 : 100.8
1963 : 100.3	1972 : 99.8	1980 : 101.0
1964 : 101.6	1973 : 98.3	1981 : 98.6
1965 : 98.2	1974 : 101.7	1982 : 100.2
1966 : 99.2	1975 : 100.2	1983 : 100.4
1967 : 101.2	1976 : 99.6	1984 : 99.7
1968 : 99.0		

In 1985, the Boston Celtics scored an average of 90 points per game.

Please continue to the next page.

SPORTS SURVEY

VAR = LOW
INFO = HIGH
ANCHOR = 100

A group of sportswriters in Boston, Massachusetts decided to examine the average points per game that the Boston Celtics scored over the past few seasons. Pro Basketball Digest showed that the average points per game scored by the Boston Celtics for the years 1960 to 1984 was:

1960 : 102.1	1969 : 100.5	1977 : 98.0
1961 : 101.1	1970 : 98.4	1978 : 98.1
1962 : 98.1	1971 : 99.4	1979 : 100.8
1963 : 100.3	1972 : 99.8	1980 : 101.0
1964 : 101.6	1973 : 98.3	1981 : 98.6
1965 : 98.2	1974 : 101.7	1982 : 100.2
1966 : 99.2	1975 : 100.2	1983 : 100.4
1967 : 101.2	1976 : 99.6	1984 : 99.7
1968 : 99.0		

In 1985, the Boston Celtics scored an average of 100 points per game.

Please continue to the next page.

SPORTS SURVEY

VAR = LOW
INFO = HIGH
ANCHOR = 110

A group of sportswriters in Boston, Massachusetts decided to examine the average points per game that the Boston Celtics scored over the past few seasons. Pro Basketball Digest showed that the average points per game scored by the Boston Celtics for the years 1960 to 1984 was:

1960 : 102.1	1969 : 100.5	1977 : 98.0
1961 : 101.1	1970 : 98.4	1978 : 98.1
1962 : 98.1	1971 : 99.4	1979 : 100.8
1963 : 100.3	1972 : 99.8	1980 : 101.0
1964 : 101.6	1973 : 98.3	1981 : 98.6
1965 : 98.2	1974 : 101.7	1982 : 100.2
1966 : 99.2	1975 : 100.2	1983 : 100.4
1967 : 101.2	1976 : 99.6	1984 : 99.7
1968 : 99.0		

In 1985, the Boston Celtics scored an average of 110 points per game.

Please continue to the next page.

SPORTS SURVEY

VAR = HIGH
INFO = HIGH
ANCHOR = 90

A group of sportswriters in Boston, Massachusetts decided to examine the average points per game that the Boston Celtics scored over the past few seasons. <u>Pro Basketball Digest</u> showed that the average points per game scored by the Boston Celtics for the years 1960 to 1984 was:

1960 : 110.0	1969 : 102.9	1977 : 99.1
1961 : 112.4	1970 : 85.0	1978 : 94.8
1962 : 92.9	1971 : 96.7	1979 : 108.8
1963 : 107.6	1972 : 90.4	1980 : 103.1
1964 : 114.8	1973 : 87.4	1981 : 96.2
1965 : 87.0	1974 : 100.3	1982 : 106.7
1966 : 93.4	1975 : 111.2	1983 : 108.1
1967 : 104.7	1976 : 86.7	1984 : 99.7
1968 : 95.3		

In 1985, the Boston Celtics scored an average of 90 points per game.

Please continue to the next page.

SPORTS SURVEY

VAR = HIGH
INFO = HIGH
ANCHOR = 100

A group of sportswriters in Boston, Massachusetts decided to examine the average points per game that the Boston Celtics scored over the past few seasons. Pro Basketball Digest showed that the average points per game scored by the Boston Celtics for the years 1960 to 1984 was:

1960 : 110.0	1969 : 102.9	1977 : 99.1
1961 : 112.4	1970 : 85.0	1978 : 94.8
1962 : 92.9	1971 : 96.7	1979 : 108.8
1963 : 107.6	1972 : 90.4	1980 : 103.1
1964 : 114.8	1973 : 87.4	1981 : 96.2
1965 : 87.0	1974 : 100.3	1982 : 106.7
1966 : 93.4	1975 : 111.2	1983 : 108.1
1967 : 104.7	1976 : 86.7	1984 : 99.7
1968 : 95.3		

In 1985, the Boston Celtics scored an average of 100 points per game.

Please continue to the next page.

SPORTS SURVEY

VAR = HIGH
INFO = HIGH
ANCHOR = 110

A group of sportswriters in Boston, Massachusetts decided to examine the average points per game that the Boston Celtics scored over the past few seasons. Pro Basketball Digest showed that the average points per game scored by the Boston Celtics for the years 1960 to 1984 was:

1960 : 110.0	1969 : 102.9	1977 : 99.1
1961 : 112.4	1970 : 85.0	1978 : 94.8
1962 : 92.9	1971 : 96.7	1979 : 108.8
1963 : 107.6	1972 : 90.4	1980 : 103.1
1964 : 114.8	1973 : 87.4	1981 : 96.2
1965 : 87.0	1974 : 100.3	1982 : 106.7
1966 : 93.4	1975 : 111.2	1983 : 108.1
1967 : 104.7	1976 : 86.7	1984 : 99.7
1968 : 95.3		

In 1985, the Boston Celtics scored an average of 110 points per game.

Please continue to the next page.

SPORTS SURVEY

Assume that you are given the task of predicting the average points per game that the Boston Celtics will score. Based on the information given on the preceding page, please estimate the probability that each of the following categories will occur (in %).

	Probability Estimate (in %)
In 1986, the Boston Celtics will average less than 90 points per game.	________
In 1986, the Boston Celtics will average between 90 to 110 points per game.	________
In 1986, the Boston Celtics will average more than 110 points per game.	________
	Total = 100%

Note: Your probability estimates should add up to about 100% because the categories include all possibilities.

Please place an X at a point on the scale below which best describes your confidence in the accuracy of your probability estimates above.

________	________	________	________	________	________
Not At All Confident	Slightly Confident	Rather Confident	Quite Confident	Very Confident	Extremely Confident

Please continue to the next page.

DECISION PROCESSES STUDY

An estimate of two more values is needed.

1) WITHOUT looking back on the previous pages of this questionnaire, please give your best estimate of the average points per game scored by the Boston Celtics basketball team:

___________ points/game

2) WITHOUT looking back on the previous pages of this questionnaire, please give your best estimate of the average amount of annual rainfall in cm. in Buenos Aires, Argentina.

____________ cm.

Thank you for participating in this study.

Appendix B

List of Abbreviations

Abbreviations

ANCHOR.	Anchor point
CONJ.	Confidence judgment
INFO.	Amount of task information
MCAT.	Mean of the categorical probability estimates
PTE.	Point estimate of the mean from memory
SCAT.	Standard deviation of the categorical probability estimates
TYPE.	Problem type
VAR.	Problem variability

VITA AUCTORIS

1959 Born Chicago, Illinois

1976 Graduated from Lakeview High School, Battle Creek, Michigan

1976 Entered University of Michigan, Ann Arbor, Michigan

1980 Graduated with a Bachelor of Science degree in Psychology from the University of Michigan, Ann Arbor, Michigan

1980 Enrolled in the Doctoral Program in Applied Social Psychology at the University of Windsor, Windsor, Ontario

1983 Graduated with a Master of Arts degree in Psychology from the University of Windsor, Windsor, Ontario

1985 Attained Doctoral Candidate Status in Applied Social Psychology Program at the University of Windsor, Windsor, Ontario