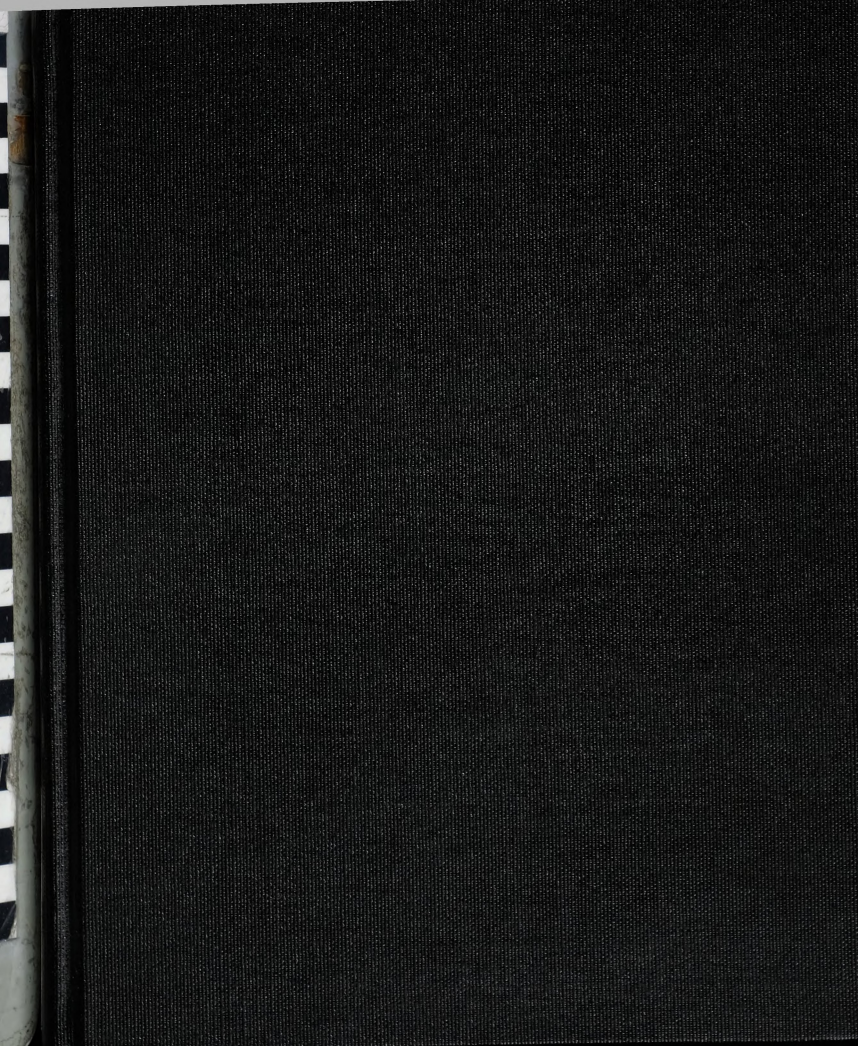
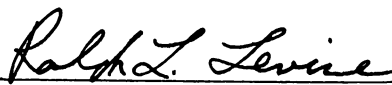






This is to certify that the
thesis entitled
A REANALYSIS OF THE BASE RATE PROBLEM THROUGH
UNDERSTANDING SUBJECTS' JUDGMENTAL REASONINGS
presented by
Wing-Shing Chan

has been accepted towards fulfillment
of the requirements for
M.A. degree in Psychology


Major professor

Date 5-31-1990

LIBRARY
Michigan State
University

PLACE IN RETURN BOX to remove this checkout from your record.
 TO AVOID FINES return on or before date due.

DATE DUE	DATE DUE	DATE DUE
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

MSU Is An Affirmative Action/Equal Opportunity Institution

c:\circ\datedue.pm3-p.1

A REANALYSIS OF THE BASE RATE PROBLEM
THROUGH UNDERSTANDING SUBJECTS' JUDGMENTAL REASONINGS

By
Wing-Shing Chan

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Psychology

1990

645-317X

ABSTRACT

A REANALYSIS OF THE BASE RATE PROBLEM
THROUGH UNDERSTANDING SUBJECTS' JUDGMENTAL REASONINGS

By

Wing-Shing Chan

This study was a response to Bruner's (1986) call for describing the process of judgment itself, in lieu of studying judgmental errors according to some mathematical or logical norms. Since the multi-modal nature of the response from base rate problems rendered traditional analysis using median problematic, qualitative methods were applied to analyze the verbal protocols of subjects' reasonings. The intuitive and probabilistic mode of judgment were delineated for classification purposes. It was discovered that the distribution of judgment were more probabilistic under problem contexts of low diagnosticity, extreme base rate and physical mechanistic environment. The same conclusions were not found for causality. Subject's sex, culture, age, major area and knowledge about the Bayes' rule showed no effect. Moreover, the probabilistic judges seemed to be less susceptible to mode shifts than the intuitive judges. It was suggested that cognitive complexity might be related to the

mode of judgment subject used. Qualitative insights also showed that subjects' apparent judgmental errors against normative rules in fact derived from cognitive and meta-cognitive skills which are vital to sound judgment in the real world.

Copyright by
WING-SHING CHAN
1990

Dedicated to
Buddhas and Bodhisattvas

ACKNOWLEDGMNETS

I wish to thank professor Ralph Levine for his heartfelt support and mind-broadening intellectual communication over the past few years. His expert advice contributed significantly to the final scientific presentation of the results and conclusions of this study. Emeritus Professor Charles Wrigley with whom I started my thesis research, has been offering me much guidance in my academic, personal and financial affairs. His challenging ideas help me strengthen my own thoughts and arguments. His sincere and useful support during my critical periods as a naive foreign student in U.S.A. deserved a great deal of merit. I wish also to thank the department chair, professor Gordon Wood who joined my thesis committee and offered excellent advice. Without their support I could not have pursued an intellectual search successfully. Their advice also furnished my critical thinking towards this research.

I am indebted to the Chinese students who participated in my experiment and gave me a chance to understand better how people make judgments. Thanks also go to my present advisor in the College of Education, professor Stephen Raudenbush whose intellectual and financial support helped

me very much during the latter period of thesis writing. He also joined my thesis examination committee. I wish to acknowledge my thanks to Erik Kvan, who initiated my critical thinking for the truth in human science while I was working as a tutor in the Chinese University of Hong Kong. I wish also to thank Chai-Liang Huang and Li-Wen Liaw, graduate students in MSU, for introducing me about Buddhism, which helped me sail through many difficulties during my study.

Last but not least, I wish to express my indebtedness to my father who died during my study in America and to my mother, brothers and sisters for my absence. Of course, my unceasing emotional support, patience and numerous help from my wife from Taiwan, Li-ming are unforgettable.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGMENTS	vi
INTRODUCTION	1
The Base Rate Problem	4
Normative Solution of the Base Rate Problem	6
THE DEVELOPMENT OF UNDERSTANDING ABOUT THE BASE RATE	
PROBLEM	8
Representativeness	9
Causality	12
Diagnosticity	14
Relevance	16
Problem solving	16
CONCEPTUAL PROBLEMS INVOLVED IN ANALYSING SUBJECT'S	
RESPONSE	19
Responding to Base Rate is Better than not	
Responding	19
Proximity to the Bayes' Norm as being Bayesian	20
Proposed Remedies	22
Obtain Additional Measures	22
Categorize the Sub-Distributions Qualitatively	22
METHODOLOGY	25
Subjects	26
Procedures	26
Materials	28
Specific Research Questions	28
Analytic Method for the Verbal Protocols	30
RESULTS	33
A Traditional Analysis of Numerical Data	33
Qualitative Categorization of Judgment Using	
Verbal Protocols	38
Reliability of coding	41
Problem Contexts and Mode of Judgment: A	
Quantitative Analysis	42
Causality and mode of judgment	45
Diagnosticity and mode of judgment	47

Extreme base rate and mode of judgment	49
Problem context (social vs. physical) and mode of judgment	51
Comparison of Consistency between Intuitive and Probabilistic Judgment	54
Effects of Culture, Sex, Age, Major and Bayes' Knowledge on Judgment	58
DISCUSSIONS	61
Qualitative Aspects of Intuitive Judgment	63
Intuitive judges do not take the information for granted	65
Intuitive judges look for more relevant information	66
Intuitive subjects supply their own knowledge and assumptions	67
Intuitive judges balance information by its relevance	69
Implications for the Debate about Human Rationality	72
Implications for the Debate about Clinical vs. Statistical Prediction	78
Cognitive Complexity as a Determinant of Judgment: A Hypothesis	80
Implications for Future Research	81
Limitations	83
APPENDIX A: Questionnaire of the Base Rate Problems	84
APPENDIX B: Two Protocol Examples	92
APPENDIX C: Classification of the Two Protocols	96
APPENDIX D: Histograms of the Numerical Judgmental Responses	102
LIST OF REFERENCES	108

LIST OF TABLES

1.	Quantitative Results based on Central Tendency . . .	34
2.	Indicators for Probabilistic Mode of Judgment . . .	40
3.	Indicators for Intuitive Mode of Judgement	41
4.	Stability and Change under Various Problem Conditions	44
5.	Effect of Causal Base Rate on Judgment	47
6.	Effect of Diagnosticity on Judgment	49
7.	Effect of Extreme Base Rate on Judgment	51
8.	Effect of Physical Context on Judgment (I)	54
9.	Effect of Physical Context on Judgment (II)	54
10.	Relative Consistency of Judgmental Mode under Various Comparisons	57
11.	Effects of Sex, Major and Knowledge about Bayes' rule on Judgment	59
12.	Correlation between age and mode of judgment	60

LIST OF FIGURES

1.	Engineer-lawyer problem	10
2.	Distribution of responses to the cab problem.	21
3.	Histogram of the subjects' responses for problem F	37
D.1	Histogram of the subjects' responses for problem A	102
D.2	Histogram of the subjects' responses for problem B	103
D.3	Histogram of frequency of subject's response for problem C	104
D.4	Histogram of the subjects' responses for problem D	105
D.5	Histogram of the subjects' responses for problem E	106
D.6	Histogram of the subjects' responses for problem F	107

INTRODUCTION

The study of error or bias in human judgment and thinking is not an invention by today's social scientists (e.g. Evans, 1989; Kahneman & Tversky, 1982; Nisbett & Ross, 1980). In a discussion about the concept of thought, Bruner (1986, p.106-107) wrote:

...It was no accident that the mathematician George Boole entitled his famous work on algebra The Laws of Thought. Thought, in this dispensation, is a normative idea, a specification of a criterion of right reason.

...it was certainly the hope of early logicians and philosophers to find some way of sorting out the chaff of unreason from the wheat of reason. And this was to be accomplished by the provision of finer and finer rules of right reason (that is, laws of logic) rather than by closer and closer description of the activity of thinking itself.

...It is curious how little psychological curiosity there was about the sources of these errors, and from the Sophists to Wurzberg one can find relatively little difference in the way they were accounted for. They were "weaknesses" in our logical processes, earlier couched in terms of weaknesses for the undistributed middle, later as "set effects" or "atmosphere effects". To put it in a word, there was no psychology of thought, only logic and a catalogue of logical errors.

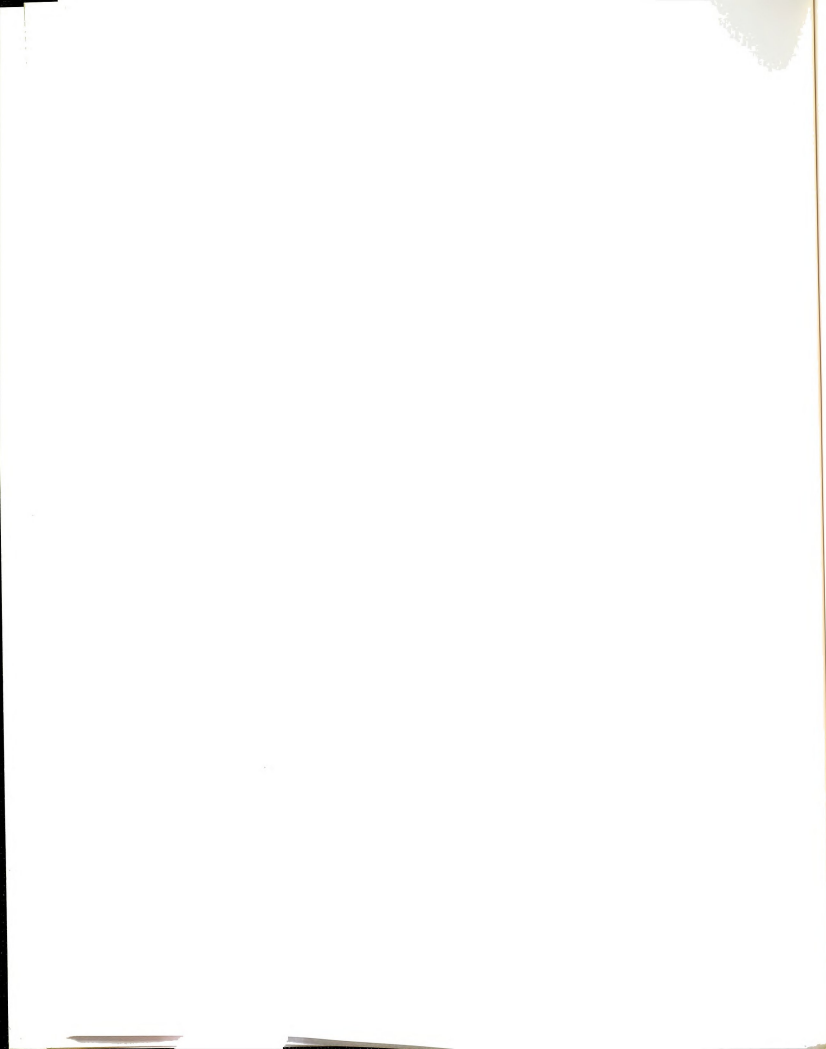
...The same case holds for the history of inference as for deduction, as with the "base rate fallacy" I discussed in chapter 6. Departure from Bayesian criteria is "fallacy", and departures as before are attributed to weakness, some to weakness induced by bias.

Bruner did not mean that the results of the studies on human judgmental errors are wrong. He actually suggested that by categorizing judgmental errors instead of describing thinking itself, researchers have not given the psychology of thought, or judgment a chance to develop.

This thesis represents a small step in response to Bruner's advocacy for describing human judgment itself instead of merely studying how judgmental errors occur. The preconditions to fulfill such a goal would at least include, as necessitated by Bruner's argument, the following two points:

1. Restraining our past tendency to study judgment according to a normative criterion and of our focus at the causation for the errors.
2. Adopting a research methodology which could maximize our chance of being able to describe the judgmental processes itself.

It is my belief that quantitative research methodology in social science research provides a vigorous and sensitive tool for detecting relations among constructs. However it is not very good at generating the most useful and interesting questions or constructs. Qualitative methodology, however, is better at describing social phenomena with detailed information which often helps

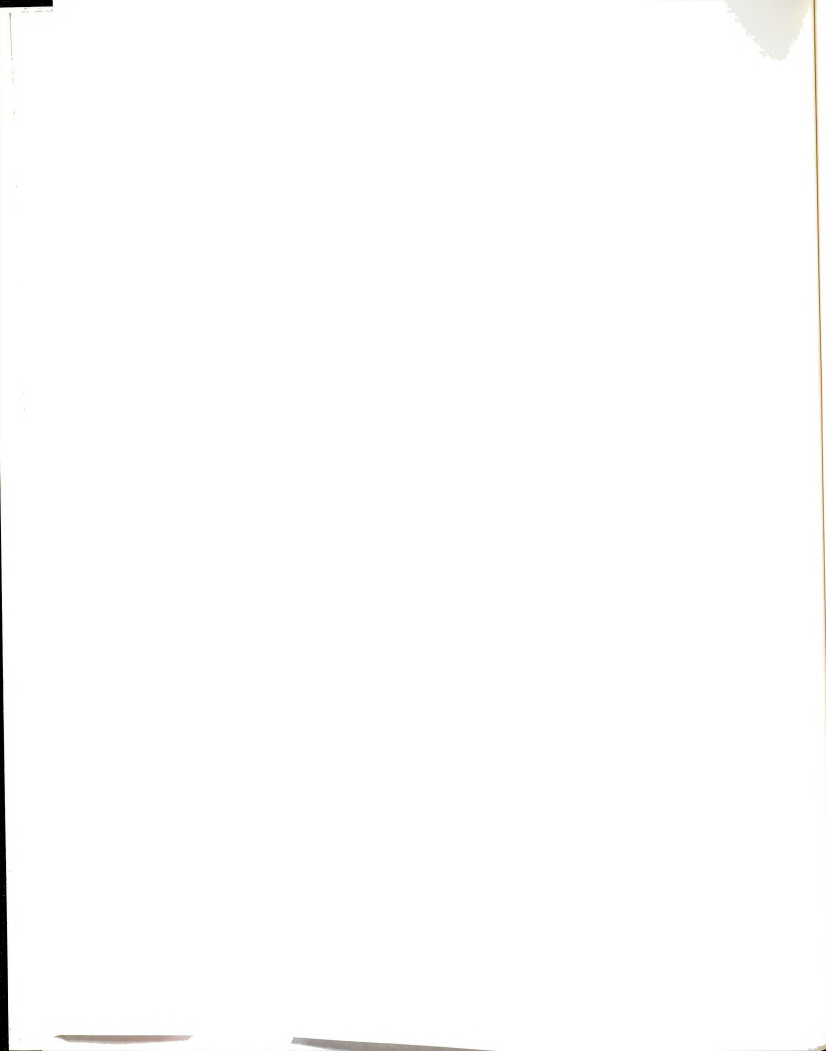


generate some insightful and useful questions and constructs. The weaknesses of qualitative research are mainly due to its poor generalization and unavailability for falsification.

One possible research methodology optimizing the advantages of both quantitative and qualitative methods is to let the latter do the job of generating ideas and the search for meaning, and let the former build a model of the resulted constructs vigorously. The present thesis adopts such an approach.

The topical research area in this thesis concerns the study of the base rate problem. Nowadays the base rate problem is one of the most intensely studied topics of inferential judgment, parallel in status to the study of syllogistic logic in deductive reasoning. In social psychological research, this area of study is often referred as the base rate fallacy (e.g. Borgida & Brekke, 1981).

The present research attempted to reanalyze the base rate problem by collecting information about subjects' reasonings. Forty Chinese students were tested individually on 7 base rate problems with theoretical interests. Verbal reasonings together with the numerical responses were used to categorize the judgment involved. The qualitative categories of judgment were used instead of the numerical



responses to analyze the effects of problem contexts such as causal base rate, non-diagnostic information, extreme base rate and physical mechanistic environment. The effect of culture, age, sex, major area and knowledge about Bayes' rule on the mode of judgment used were also studied. The stability of the two mode of judgment against experimental treatments were also investigated using appropriate statistical tests. Qualitative aspects of the intuitive judgment was described and the implications to the debate about human rationality was also discussed.

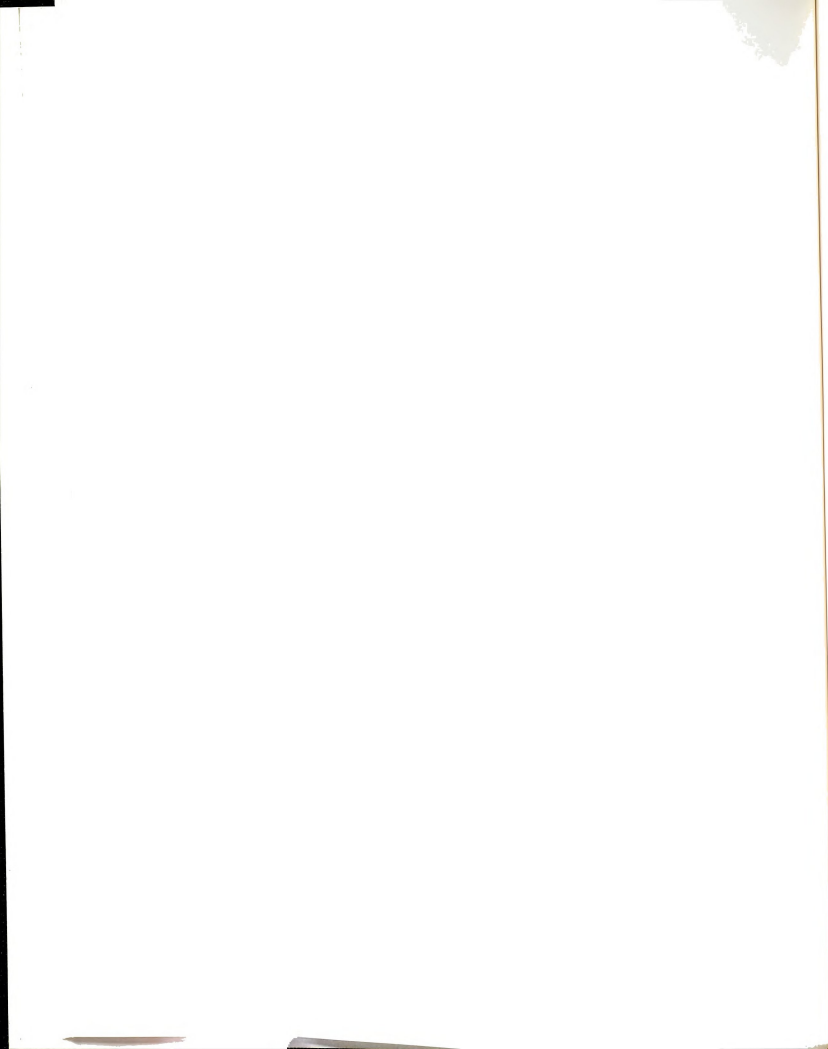
The Base Rate Problem

The base rate problem normally experimented in psychology is in fact a mathematical Bayes' problem with two outcomes. The following is an example taken from Tversky and Kahneman (1980), commonly referred to as the "cab" problem:

A car was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

(a) 85% of the cabs in the city are Green and 15% are Blue.

(b) a witness identified the Cab as Blue. The Court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time.



What is the probability that the cab involved in the accident was Blue rather than Green?

Analytically, this problem is constructed according to the mathematical Bayes' formula with two mutually exclusive elements. In this case the outcome is either a blue cab or a green cab. The subject is asked to determine the probability for a given outcome.

Another common element of the base rate problems is the base rate information. It is the statistical probability for an outcome given no further information about an event. For example the base rate information is 85% for green cab and 15% for blue cab. Since the outcomes are mutually exclusive, the addition of their probability must be equal to unity.

The third important element in these problems is the "diagnostic information", so termed for our convenience when discussing research on the base rate problem. This information gives us specific information about the occurrence of an event in addition to the base rate information. For this example the diagnostic information is the reliability information of the witness. Using Bayes' formula, a normative solution can be computed.

Normative Solution of the Base Rate Problem

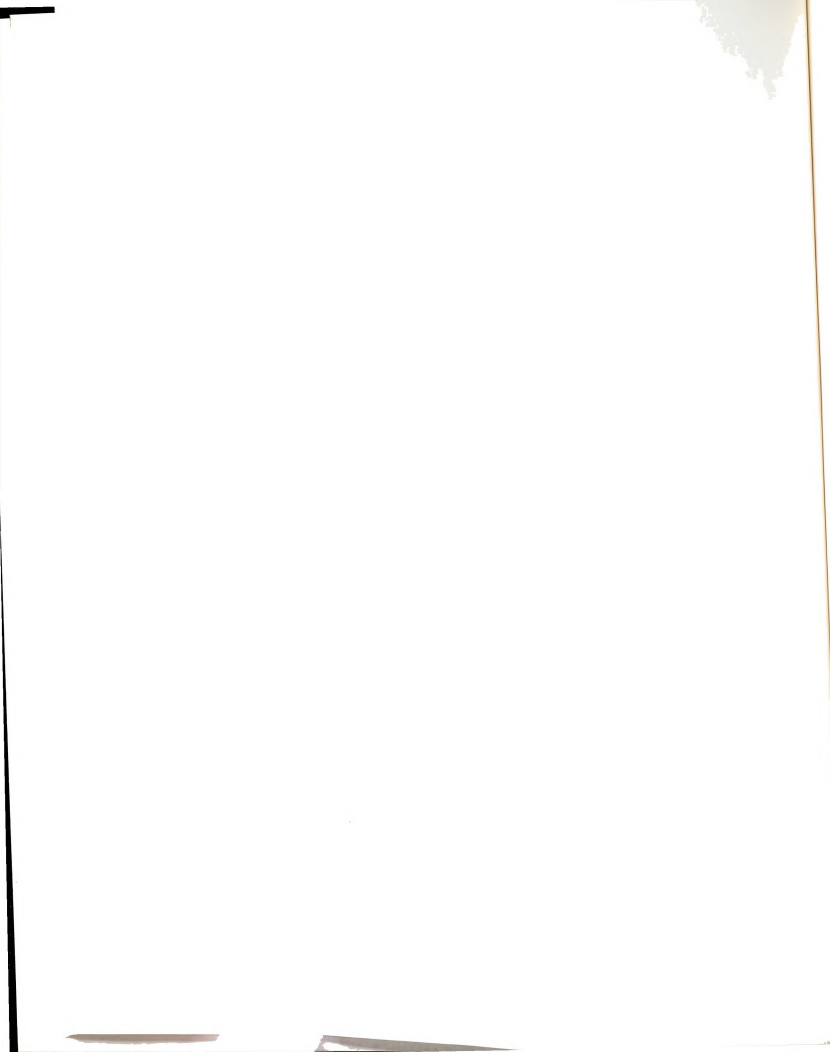
Let us explain how a Bayesian optimum can be computed for the base rate problem when both the base rate information and the diagnostic information are expressed in numbers.

Let $P(C)$ be the probability of occurrence for the outcome category C . $P(C^{\sim})$ will be the probability for the mutually exclusive event of C , called $C^{\sim}$. Following the fundamental axiom of mathematical probability theory, $P(C) + P(C^{\sim}) = 1$. In our case, the base rate information is given by $P(C)$ and $P(C^{\sim})$. These probabilities are sometimes referred as the prior probabilities.

The probability that category C has occurred given the diagnostic information D is the probability that takes into consideration of the diagnostic information as well as the base rate information. This probability is referred in mathematics as $P(C/D)$, read as probability of C given D . $P(C/D)$ and $P(C^{\sim}/D)$ are sometimes referred as the posterior probabilities. The Bayes' formula for this probability is as follows:

$$P(C/D) = P(D/C)P(C) / (P(D/C)P(C) + P(D/C^{\sim})P(C^{\sim}))$$

For the previous cab problem, the probability that the blue cab involved in the accident given the fact that the witness has identified the cab as blue with a certainty of 80% can be computed by following the above equation.

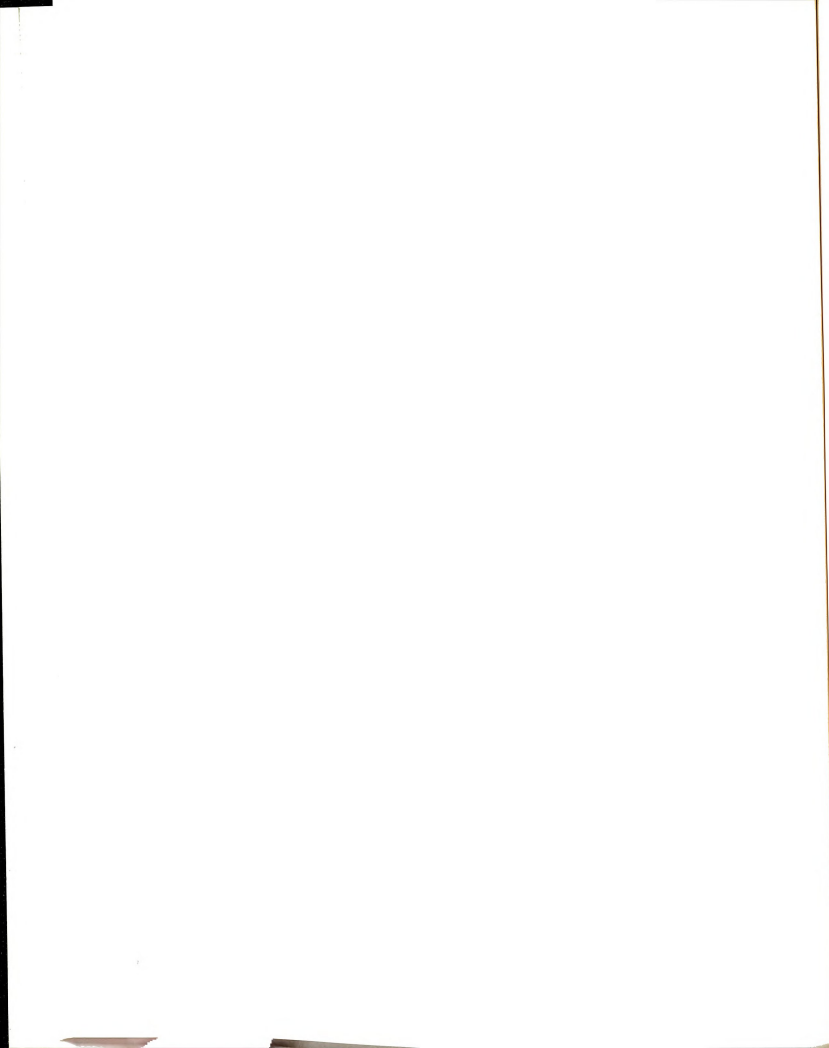


The required probability is $P(\text{blue cab}/\text{identified as blue})$, the base rate information is $P(\text{blue cab})=0.15$ and $P(\text{green cab})=0.85$. $P(\text{identified as blue}/\text{blue cab})$ is the witness's identification ability, is therefore 0.80. $P(\text{identified as blue}/\text{green cab})$ is the error rate in identification of the witness, and is thus 0.20, assuming that the error rate is the same when identifying the blue cab or the green cab. Accordingly,

$$\begin{aligned} P(\text{Blue cab}/\text{identified as blue}) \\ &= 0.80 \cdot 0.15 / (0.80 \cdot 0.15 + 0.20 \cdot 0.85) \\ &= 0.414 \end{aligned}$$

The required answer is thus 41.4%

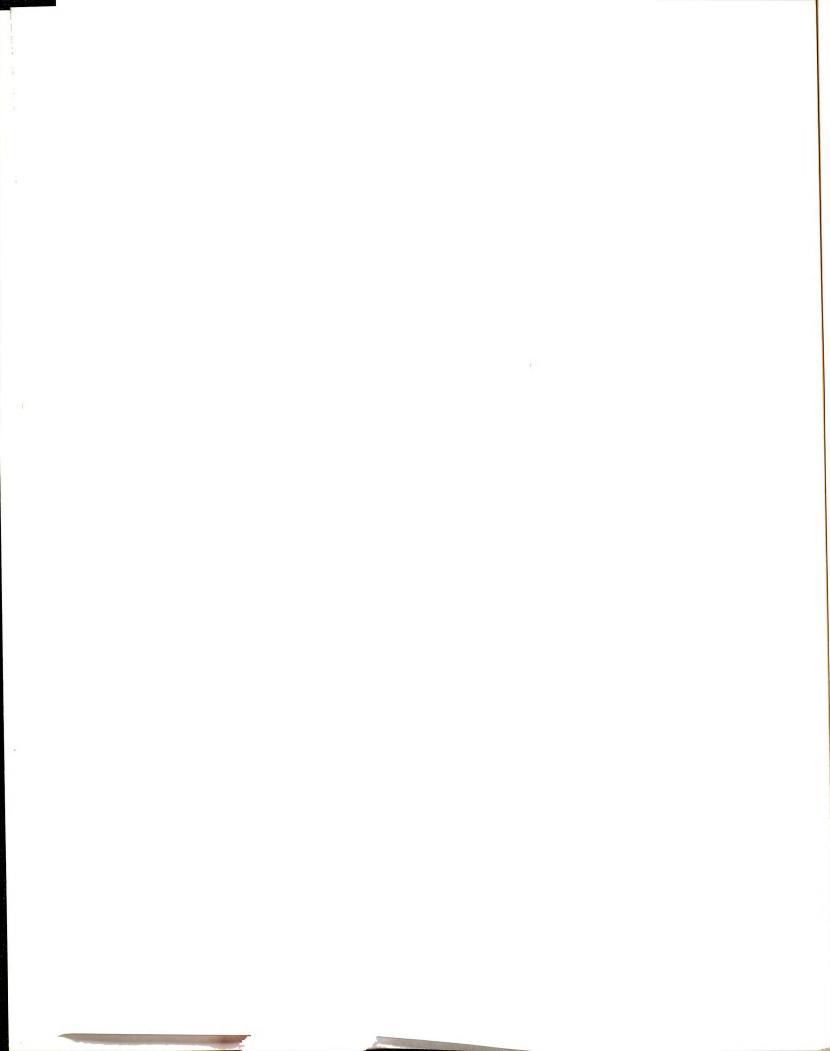
Generally, the base rate problem is used for investigating the use of base rate in judgment or decision making, as well as to compare people's judgment with the optimal judgment accorded by the Bayes' theorem. Researchers are also interested in studying the various factors which would make people more prone to making a Bayesian optimal judgment and the factors which would affect how subjects combine the base rate information and the diagnostic information while making judgment.



THE DEVELOPMENT OF UNDERSTANDING ABOUT THE BASE
RATE PROBLEM

The attempts to investigate and explain people's non-Bayesian behavior began when Kahneman and Tversky (1973) called our attention to the pitfalls of human judgment against the mathematical norm. Explanations were sought to explain why people commit errors in judgment. Later on as further research (c.f. Borgida and Brekke, 1981) found that people do use the base rates and/or give answers close to the Bayesian optimum under certain experimental manipulations, explanations were refined to contain both Bayesian and non-Bayesian behaviors and to state the conditions under which each behavior occur.

It is beneficial to review the literatures chronologically to understand why certain explanation forms have come and gone. The following review attempts to highlight the major development of explanations about the base rate problem. It does not pretend to contain every base rate studies nor every technical subtlety affecting the use of base rate. For a longer review of other details, see Borgida and Brekke (1981).

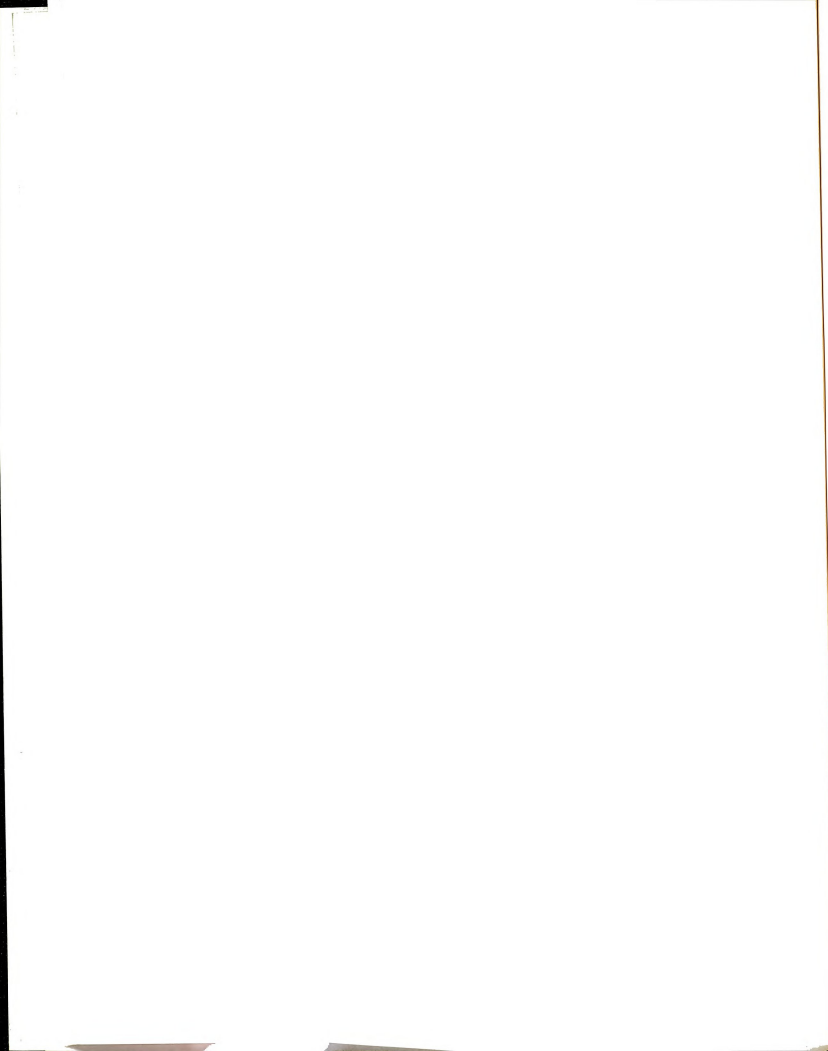


Representativeness

In the now classic engineer-lawyer problem (Listed as problem C in appendix A) designed by Kahneman and Tversky (1973), it was discovered that people seemed to judge a randomly selected personality description according to how well the description represented a typical engineer's characteristics irrespective of the ratio of the number of engineer's versus lawyer's descriptions in the sample.

Each of the subjects were given five personality descriptions consecutively and were told that each descriptions were randomly selected from 100 descriptions of a group of people consisted only of engineers and lawyers. The subjects were also divided into two groups. In one group they were told that the initial ratio of the number of engineers versus lawyers was 70:30 and the other group 30:70. The subjects were requested to judge the probability of each of the five personality descriptions as being belonged to an engineer.

If the subjects were sensitive to the prior probabilities, the estimated probabilities from the two groups with different priors should be different. However estimations from the two groups almost lie on a 45 degree straight line on a cartesian plane away from the normative Bayesian curve (see figure 1). The results showed that the



two groups' answers were almost identical, not sensitive to the difference in prior probabilities and of course not consistent with the predictions from Bayes' theorem.

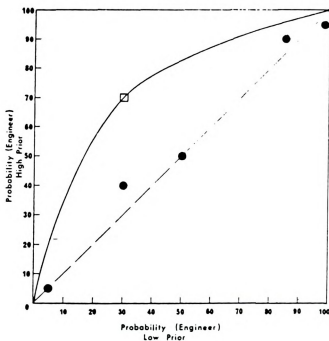
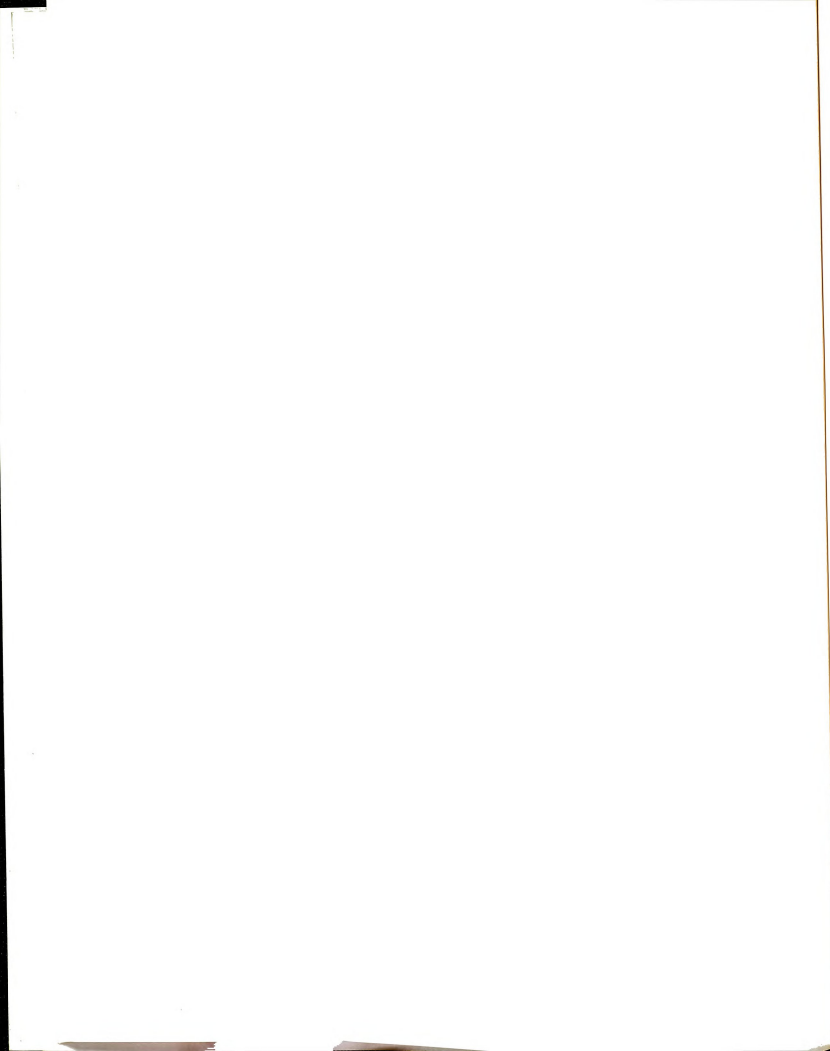


Figure 1: Engineer-lawyer problem. Median judged probability (engineer) for five descriptions and for the null description (square symbol) under high and low prior probabilities. (The curved line displays the correct relation according to Bayes' rule.) (Source: Kahneman, Slovic and Tversky, 1982, p.55.)

When the personality description looked like a typical engineer's characteristic, subjects responded with a high median probability, i.e. 90 - 100%. But when the description

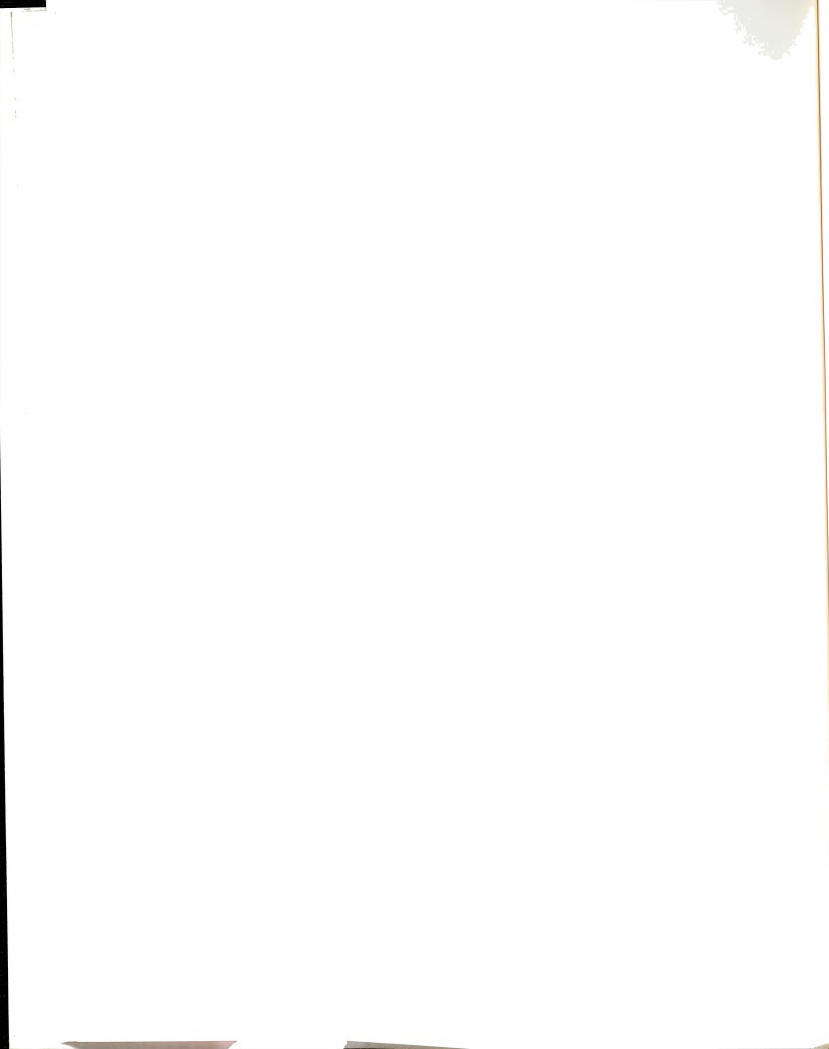


offered no specific information, subjects respond with about 50% certainty. Subjects' degree of certainty seemed to vary with the extent of which the description looked like a typical engineer.

Therefore Kahneman and Tversky (1982) concluded:

Given specific evidence ..., the outcomes under consideration ... can be ordered by the degree to which they are representative of that evidence. The thesis of this paper is that people predict by representativeness, that is, they select or order outcomes by the degree to which the outcomes represent the essential features of the evidence. (p.48)

Basically Kahneman and Tversky (1973) tried to show that people's judgement were not affected by base rates but only followed the representativeness of the diagnostic information. The only case under which Kahneman and Tversky's (1973) subjects followed base rates was when no diagnostic information of any kind was given. However, we cannot thereby say that their subjects showed some signs of using base rates because the base rate problem without diagnostic information should not be considered as a base rate problem. (see our definitions for base rate problem in Chapter 1) In short, the early research done by Kahneman and Tversky (1973) and others (e.g. Hammerton, 1973; Lyon and Slovic, 1976) showed a period of non-Bayesian behavior, nonuse of base rates and misjudgment.



Causality

Later research with varying experimental manipulations began to show that people do use base rate under certain conditions. Ajzen (1977) was first to point out:

In contrast to previous research, it was found that people's predictions were strongly influenced by base rate information but only to the extent that the base rates had causal implications for the criterion. When the base rates did not have such causal implications, they were largely neglected in favor of diagnostic information. (p.303)

For example, in one of Ajzen's (1977) factorial experiment, subjects were requested to judge from a personality outline of a factitious person, the probability to pass a final examination. Included within the study were two types of base rate conditions. In the causal base rate condition, subjects were given the information that 75% (or 25%) of the students had passed an examination of the same course two years ago. For the noncausal base rate condition, subjects were told that a certain educational psychologist interviewed some of the same students who passed the exam two years before. 75% (or 25%) of his sample passed the exam.

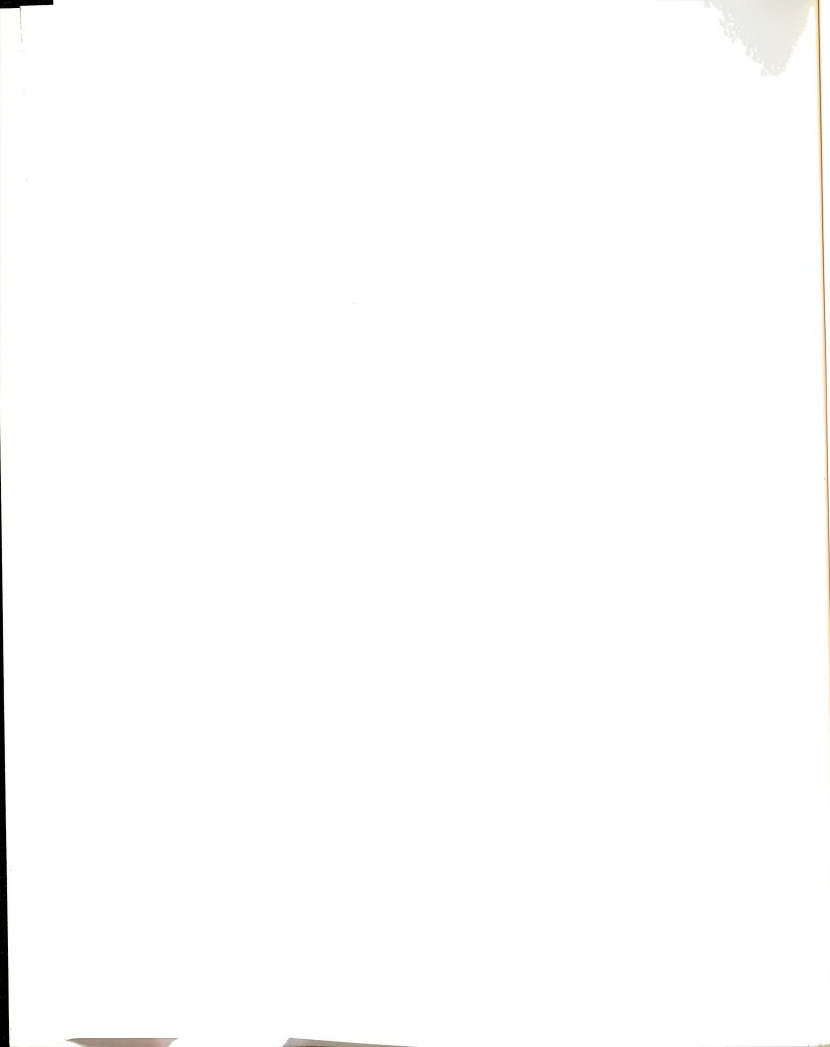
A post-hoc test was used to confirm that the 25% passing rate was perceived by subjects as a significantly more difficult exam than the 75% one. Ajzen (1977, p.304) thereby thought that the inferred exam's difficulty level

"does have a causal effect on a given student's success or failure".

The results showed a significant interaction between the base rate of success (75% vs. 25%) with the type of base rate (causal vs. non-causal) as well as a main effect on the base rate of success. And the causal base rate had a stronger effect on prediction of exam success than the noncausal base rate.

All the results in Ajzen's experiment taken together indicated that different base rates do had different effects on people's judgment. The effect of base rate was largest when the base rate had a causal implication for the diagnostic evidence. When the base rate was noncausal, the effect of base rate was minimal and people would judge mainly by means of the diagnostic information.

Ajzen's (1973) results had given us a more precise picture about people's use of base rate than the early Kahneman and Tverskys' (1973). Unlike the latter authors' claim, Ajzen discovered that people do use base rate in their judgment, at least when the base rate information has a causal implication. However, Ajzen's results were consistent with the representativeness proposal when the base rate was a noncausal one.

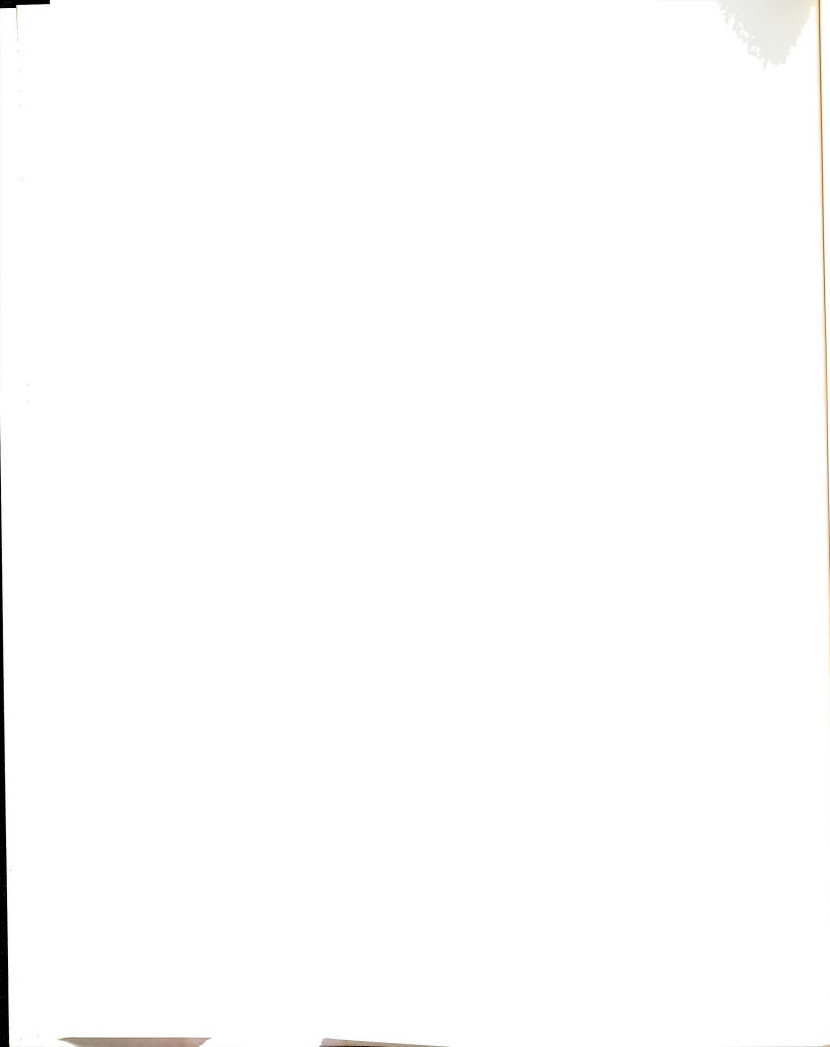


Diagnosticity

Since the discovery of the base rate fallacy, people start to investigate the effect of different types of base rates by varying the quality of the base rate, such as causal vs. noncausal one. It was discovered by Ajzen (1977) that people are sensitive to a base rate with a causal implication towards the diagnostic information. Very naturally, the next step would be to investigate the diagnostic information by varying this variable in order to study its effect on people's use of base rate. This task was taken up by Ginosar and Trope (1980). They pointed out that causality alone does not determine the use of base rate, the validity of the diagnostic information also affects the use of base rate.

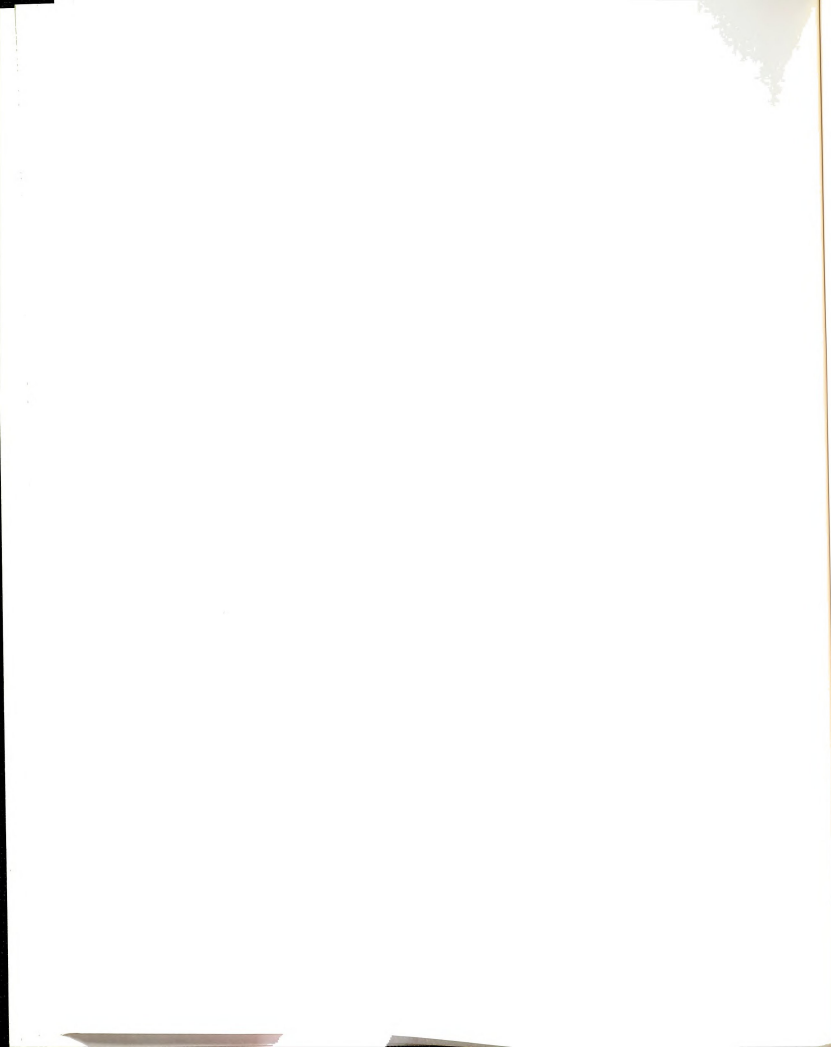
Ginosar and Trope (1980) restudied the engineer-lawyer problem by adding a diagnostic condition with inconsistent information. This condition contained information with implications for both engineer and lawyer. Subjects' median responses then varied in direct proportion to the variation in base rates. Parallel results were also demonstrated with their 'field-of-study' problem.

Results showed that, in addition to causality, the validity or diagnosticity of the diagnostic information also plays an important part in determining the use of base rate.



Similar results were obtained in other studies by varying explicitly the degree of diagnosticity or accuracy of individuating information (Fischhoff & Bar-Hillel, 1984; Hinsz, Tindal, Nagao, Davis & Robertson, 1986). In addition, when some unrelated information was added to a diagnostic information, the diagnosticity would be diluted rapidly (Nisbett, Zuckier & Lemley, 1981).

All the studies cited in this section demonstrate that diagnosticity of information can influence the way people make judgment. People tend to rely on diagnostic information when the diagnosticity is high, and rely on the base rate when the diagnosticity is made minimal. These findings are consistent with the early Kahneman and Tverskys' claim of representativeness except that there are proven conditions under which people consistently made more use of base rate. The diagnosticity explanation can join with the causality explanation to co-determine probability judgment and the use of base rate. A later study by Hinsz et al. (1981) indeed demonstrated that although causal nature of the base rate factors had a significant effect on subject's probability judgment, it was relatively minor in comparison with the impact of the accuracy of the source information, or diagnosticity.



Relevance

Researching on the side of the base rate, Bar-Hillel (1980) argued that relevance, not causality per se, determines the use of base rate. Using a modified cab problem, Bar-Hillel showed that subjects' responses came closer to the Bayesian optimum when the base rates of cabs at the region closer to the neighborhood of the accident was additionally stated. Bar-Hillel argued that such sub-group of base rate becomes more relevant and will be integrated by subjects together with other information.

In short, we can see that in order to determine the use of base rate or probability judgment, relevance is important on the side of base rate factor, and diagnosticity is vital on the side of the diagnostic information.

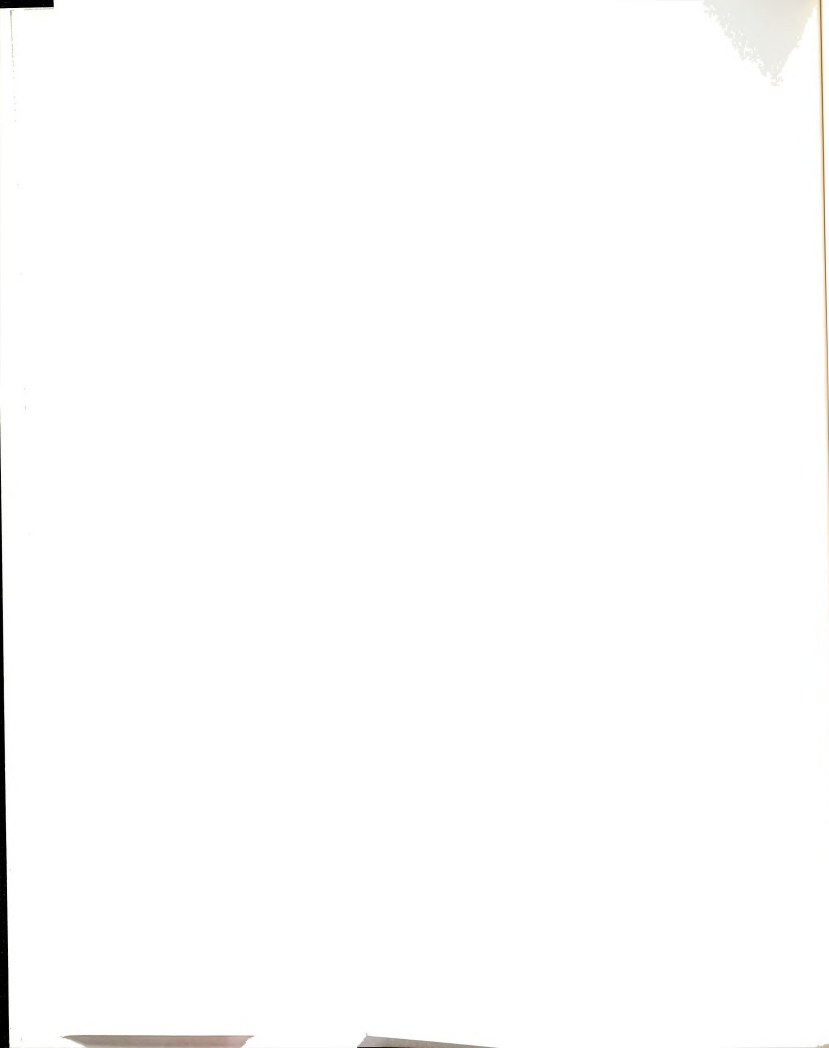
Problem solving

Through a series of experimental manipulations, Ginosar and Trope (1987) had successfully demonstrated that people's judgment under uncertainty do vary under a number of new set of conditions. These researchers argued that judgment under uncertainty can be explained parsimoniously by the "problem solving" approach.

For example, probability judgment was found to depend on prior problem diagnosticity. When prior problems had non-

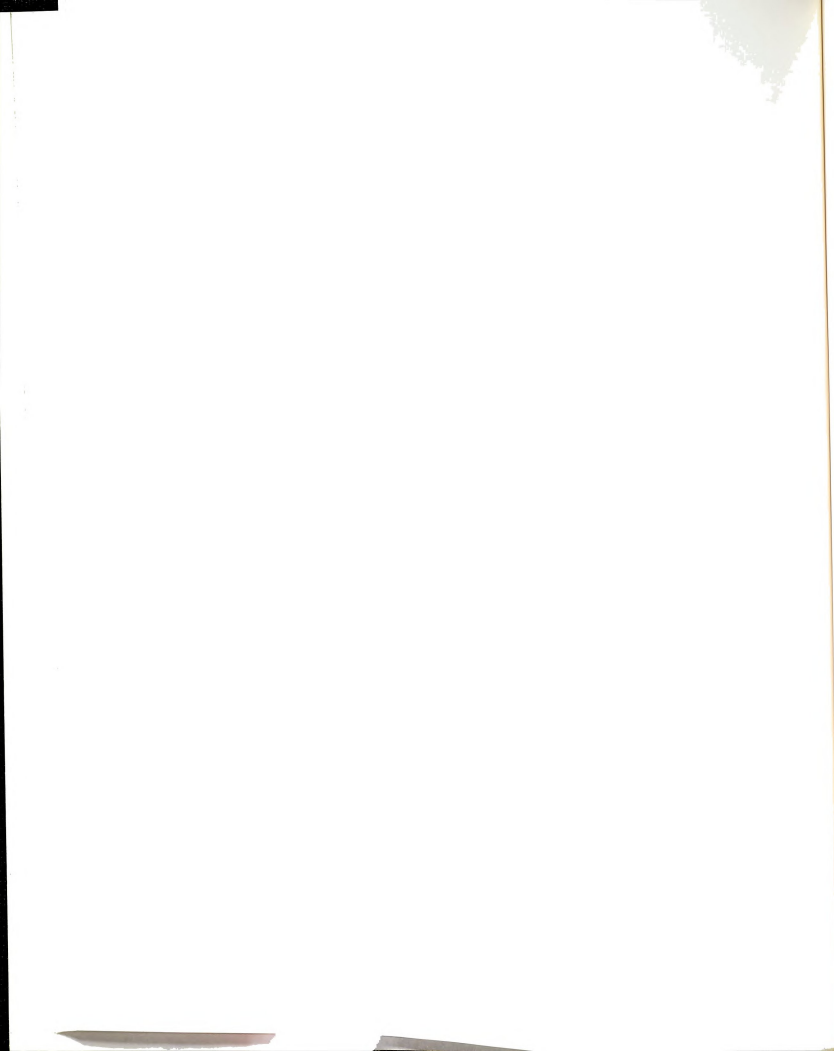
diagnostic conditions, the judgment of the problem that followed would exhibit a higher use of base rate than when preceded by a problem with diagnostic condition. Ginosar and Trope offered the explanation as 'prior activation of inferential rules'. In a second experiment, the original engineer-lawyer problem was listed in a sentence by sentence format. The change in mean probability judgment was explained as 'concurrent activation of inferential rules'. In the third experiment, probability judgment was found to vary according to whether the correct category (in a Bayesian sense) was initially given to the subject or not. Ginosar and Trope again related this effect to the goal-directedness in problem solving theories. By decreasing the source reliability of the diagnostic information in the engineer-lawyer problem and converting the cab problem to resemble drawing marbles, significant decrease in probability of judgment were observed. These phenomena were explained as restrictions on the application of the representativeness rule and enhancement of the applicability of the sampling rule respectively.

Ginosar and Trope were the first researchers who attempted to uphold a coherent and consistent theoretical framework (i.e. problem solving approach) to explain the various experimental results in base rate research. Their



effort still represents by now the most encompassing theoretical work in this field.

In conclusion, our review of the major literature confirms Bruner's claim that almost no research effort is directed to the study of thinking or judgment by themselves. All the research cited in this chapter is only concerned with whether subjects are making correct judgment according to the normative Bayes' theorem.



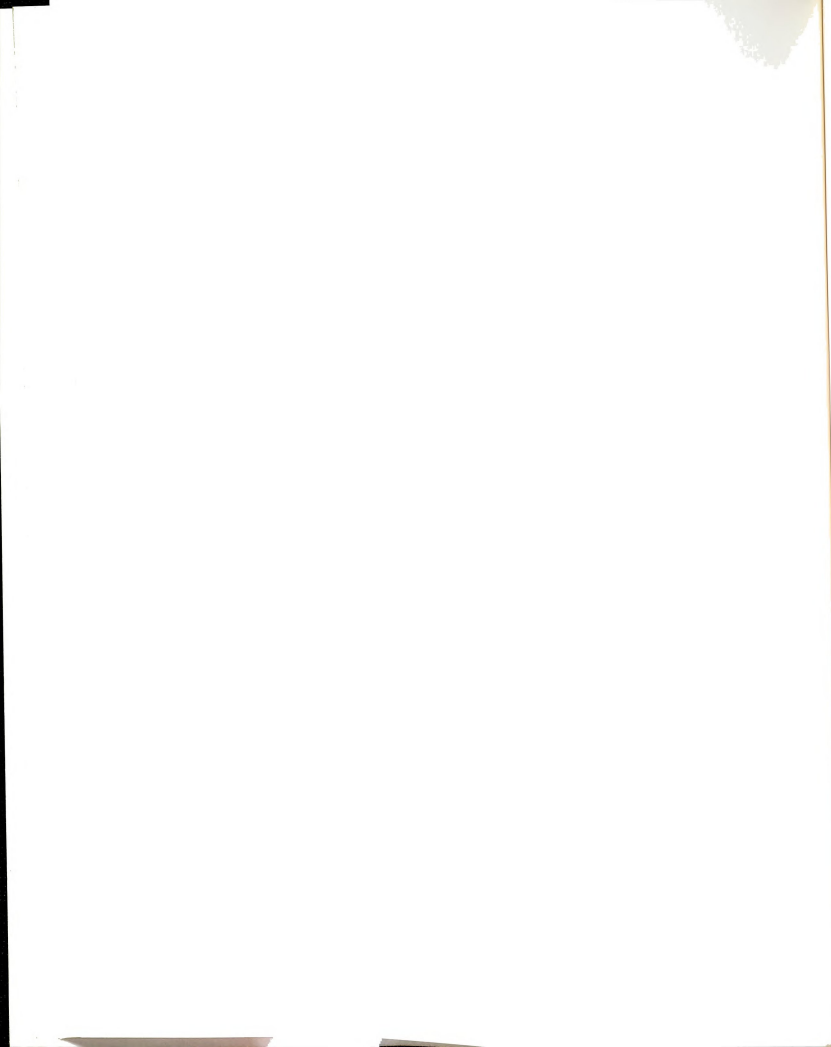
CONCEPTUAL PROBLEMS INVOLVED IN ANALYSING
SUBJECT'S RESPONSE

As usual, researchers generally computed and presented the mean or the median judgmental responses (e.g. Bar-Hillel, 1980; Ajzen 1977, Ginosar & Trope, 1987, etc.). The central tendency is either compared to the Bayesian Optimal value or compared against treatments to obtain causal relationship between treatments and the mean judgmental responses.

However, the above common data analytic procedures and interpretation contain, I think, at least two unjustified beliefs, intermixed with conceptual and technical difficulties. They are listed as follows:

Responding to Base Rate is Better than not Responding

One of the designs illustrating this belief is set up via a ANOVA (e.g. Ajzen, 1977). A base rate problem is given to two groups of subjects, with base rate being different in the two groups. When the mean responses of probability judgment are significantly different between the two groups, researchers obtain evidence that subjects are sensitive to the magnitude of base rate. Since a Bayesian



scenario involves the mathematical weighing of the base rate information and the diagnostic information, the proven sensitivity to base rate level is generally construed by researchers as being a better judgment, at least better than subjects who seem to concern about the diagnostic information only (see Kahneman, Slovic & Tversky 1982).

However, a priori speaking, using base rate does not guarantee the final answer to be Bayesian equivalent, because, in actual practice, there are numerous ways of using the base rate. The outcome is a priori unpredictable with respect to Bayesian optimum.

Proximity to the Bayes' Norm as being Bayesian

A mean or median response close to the Bayesian optimal value is regarded as the better Bayesian judgment. There are two problems with regard to this belief. First, it is a priori possible for some non-Bayesian behaviors to get answer close to the Bayes' optimal value. Second, distribution of response in base rate research are often bimodal or multi-modal (e.g. see Figure 2), using mean or median as central tendency are not a fair measure of what subjects are doing.

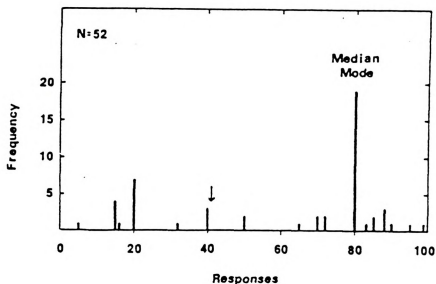


Figure 2: Distribution of responses to the cab problem.. The arrow indicates the correct Bayesian estimate. (Source: Bar-Hillel, 1980)

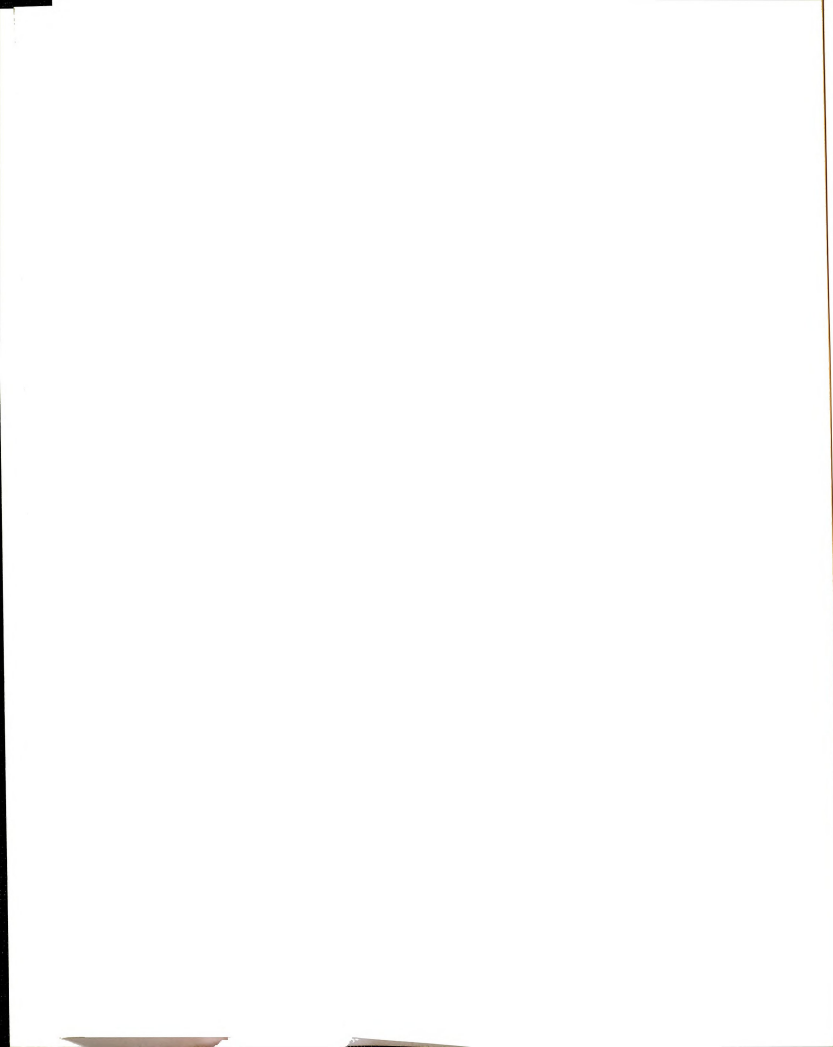
In fact, the underlying generative mechanisms within each sub-distribution might be different. Some might be close to Bayesian behavior while some might not. Therefore the central tendency measure using mean or median actually mis-represent the underlying mechanisms with regard to Bayesian optimum. As a consequence, the distance between the central tendency measure and the Bayesian optimal value may be rendered non-interpretable.

Proposed Remedies

In relation to the conceptual problems in analysing subjects' judgmental response, the following two remedies are proposed:

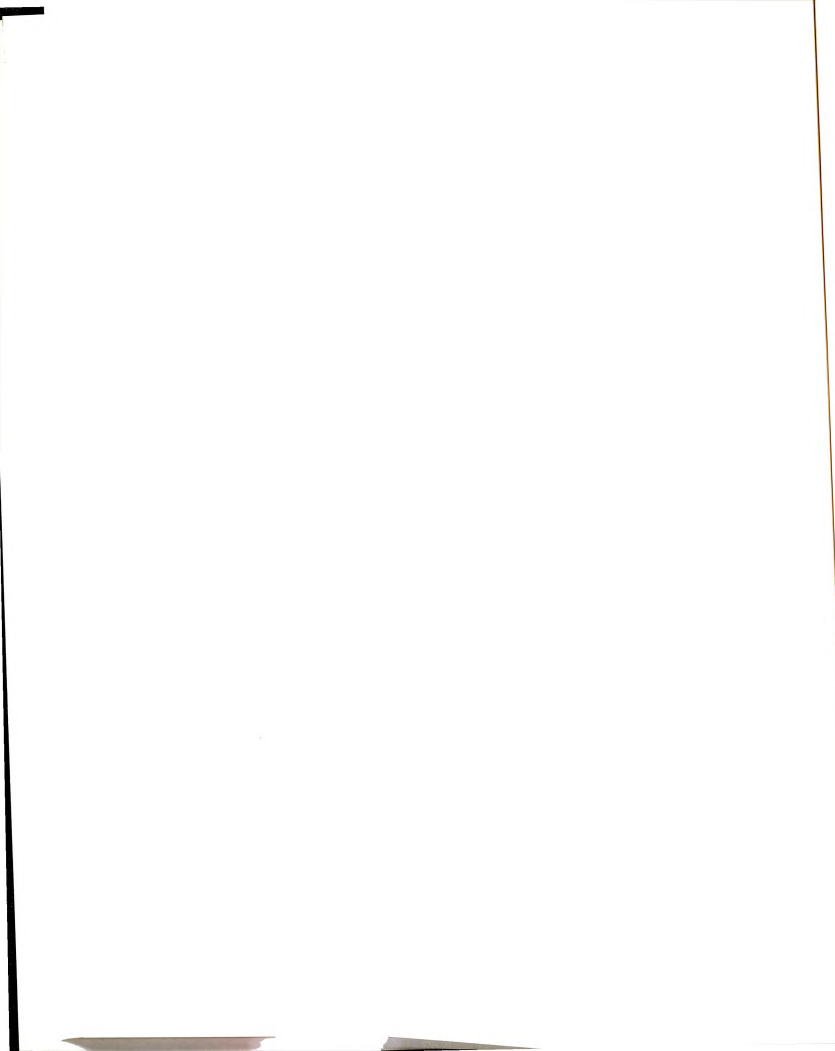
Obtain Additional Measures. A particular defect in former base rate studies is that, by focusing on subjects' final responses, one loses sight of the underlying generative processes or mechanisms of the subjects' solutions. Accordingly one loses the fundamental grounding in deciding whether the subject's behavior is Bayesian or non-Bayesian. It is suggested that one can use post-experimental interview or thinking-aloud procedures (for trained subjects) to investigate the explanations, reasonings and conscious processes of thinking (Newell and Simon; 1972). By contrasting the final judgmental responses with the reasoning of subjects, a better measure of whether the subjects' behavior is Bayesian or non-Bayesian can be obtained. Even though it turns out to be non-Bayesian, the results would still open a new horizon of research into the structure of people's reasoning in addition to obtaining their numerical probability judgmental responses.

Categorize the Sub-Distributions Qualitatively. A simple way to solve the multi-modal distribution problem



might be to assign the separate sub-distributions integer values and use chi-square to test the change in the sizes of the sub-distribution under different experimental conditions with theoretically important implications. However, there might be borderline cases in which it may be difficult to classify by numerical magnitudes of response to which sub-distribution they belong, or whether these cases should be considered as a separate meaningful subgroup. Independently obtained qualitative categorizations of the verbal protocols obtained from subjects as recommended in last section might be useful in solving this classifying problem. Some useful qualitative classifying techniques can be adopted from 'phenomenography' (Marton, 1981) or from 'grounded theory' (Strauss, 1987).

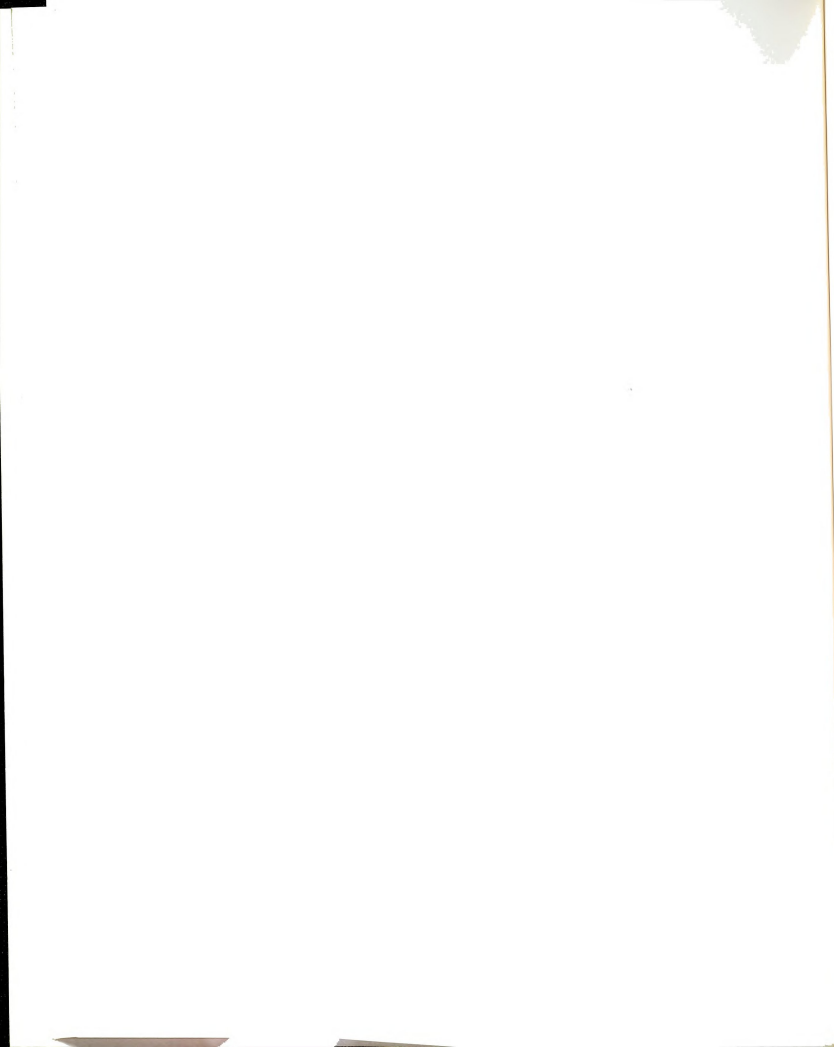
The common features of these classifying techniques involve careful coding of individual verbal protocols. The protocols are brought together into groups on the basis of similarity and the groups can be compared to each other. A higher order of meanings which emerged are combined to form the categories of descriptions. The distinctive feature of this method is that 'the analysis is dialectical in the sense that bringing the quotes together develops the meaning of the category, while at the same time the evolving meaning determined which of the categories are included or omitted' (Marton & Saljo, 1984, p.55). A more detailed description of



this technique would be discussed in the "methodology" chapter of this thesis.

Readers might question the subjectiveness of the coding process involved. This can be answered by noting that qualitative categories are constructed from data to understand the phenomena, they are by no means final and are subjected to modification or synthesis when provided with more data or when the research focus shifts (see Strauss, 1987).

As long as the schemes of classification and coding are carefully and explicitly laid out, the categorization processes can be repeated by independent judge to check the reliability of coding using this schemes. Repeated research can provide information to validate or reject the usefulness of the kind of coding in relation to specific research purposes (Strauss, 1987).

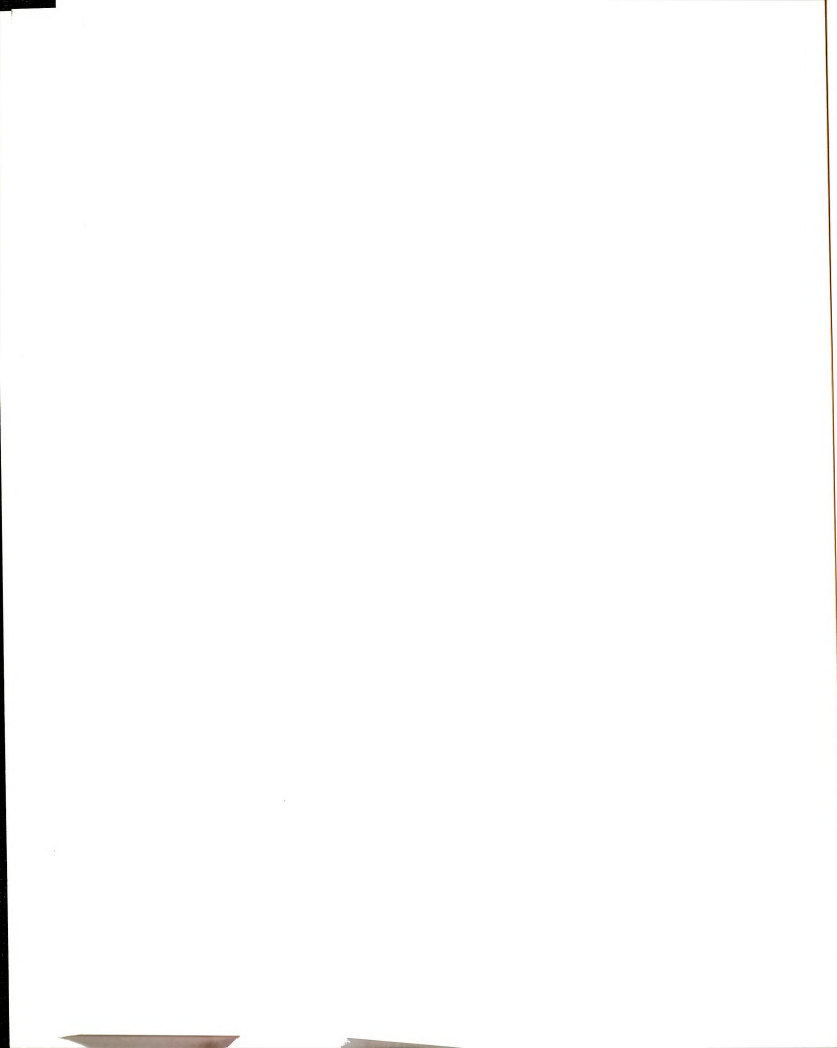


METHODOLOGY

The methodology adopted in this work is a response to Bruner's call for a study of judgment itself as well as a response to the conceptual and technical problems in base rate research as discussed in the last section.

Our attempt is to reanalyze the base rate problem through understanding subjects' judgmental reasoning. Whether subjects' responses comply with the normative rule is not the primary interest. The investigation of subjects' judgmental reasonings is achieved by collecting verbal data in addition to numerical responses. Subjects were asked to think aloud in judging several base rate problems of theoretical interest. They were then interviewed by the experimenter with regard to their judgmental reasonings.

Qualitative techniques were applied to the analysis of the verbal data. Useful and meaningful constructs of the judging process were delineated for subsequent quantitative analysis. Qualitative insights would also be observed.



Subjects

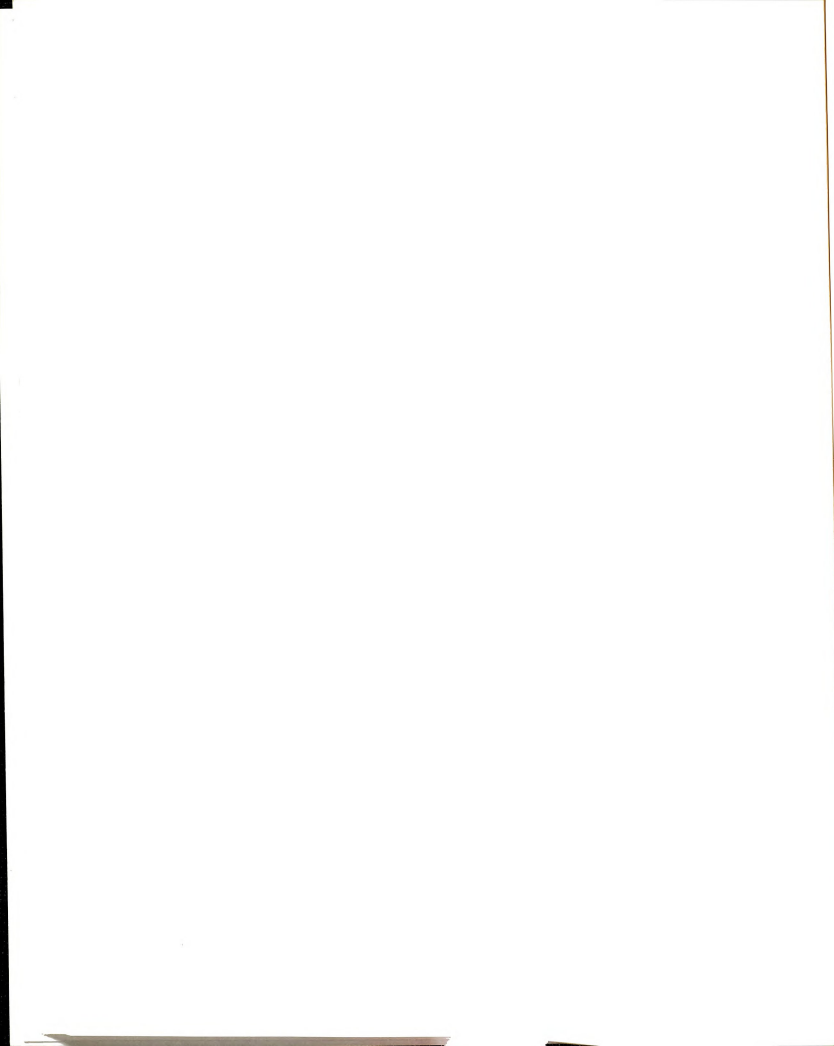
The subjects were 40 volunteers recruited from Chinese students on Michigan State University campus, 20 males and 20 females. For each gender, half of them were chosen with backgrounds in arts or social sciences and the remaining half with a major in science or engineering.

Procedures

Each subject was tested individually with the experimenter by his side. The testing session generally involved 40 minutes to 1 hour. The questionnaire was typed in English while all colloquial interchanges were mainly in Chinese. Mandarin was used for students who came from Mainland China or Taiwan. For Hong Kong students, the Cantonese dialect was used.

First, the subject was allowed to read the first page of the questionnaire containing the instructions. (The whole questionnaire is listed in the appendix of the thesis) Then the interviewer would give the following statements:

This research is to study how ordinary people make judgment on certain everyday affairs. You would have seven problems to do. In all of these problems, there are no absolute answers of any kind. You don't have to worry whether your answer is right or wrong. Therefore you can use your own methods to make what appears to you the best judgment. In the beginning you would use a method called thinking aloud method. That means you try to say what you are thinking about when you are thinking over the problem. Just tell us what comes



to your mind and we would tape-record it. After you have used this method to finish all of the problems, I would interview with you and ask you what your reasoning is and how you come to the answer. The questionnaire generally takes 20 minutes to complete and the interview session would last for a further 20 minutes. There is no time limit in completing the questionnaire. You can do it at your own pace. If you have any questions about the meaning of the wordings in the questionnaire, feel free to ask me. Do you have any questions? ... If not, you can begin. "

While the subject was thinking aloud, the interviewer put down the main arguments or reasoning processes spoken out by the subject. After the thinking aloud procedure, the interviewer proceed to interview the subject on each problem about the subject's reasoning processes.

To be economical while still remaining accurate, the verbal protocols from the think aloud and research interview sessions were written simultaneously during the experiment. The protocols were written at sufficient detail to capture the main reasonings of the subject. Examples of the protocols can be found in the appendix. Tape recordings were only referred to whenever there were some problems in understanding the protocol for the subject.

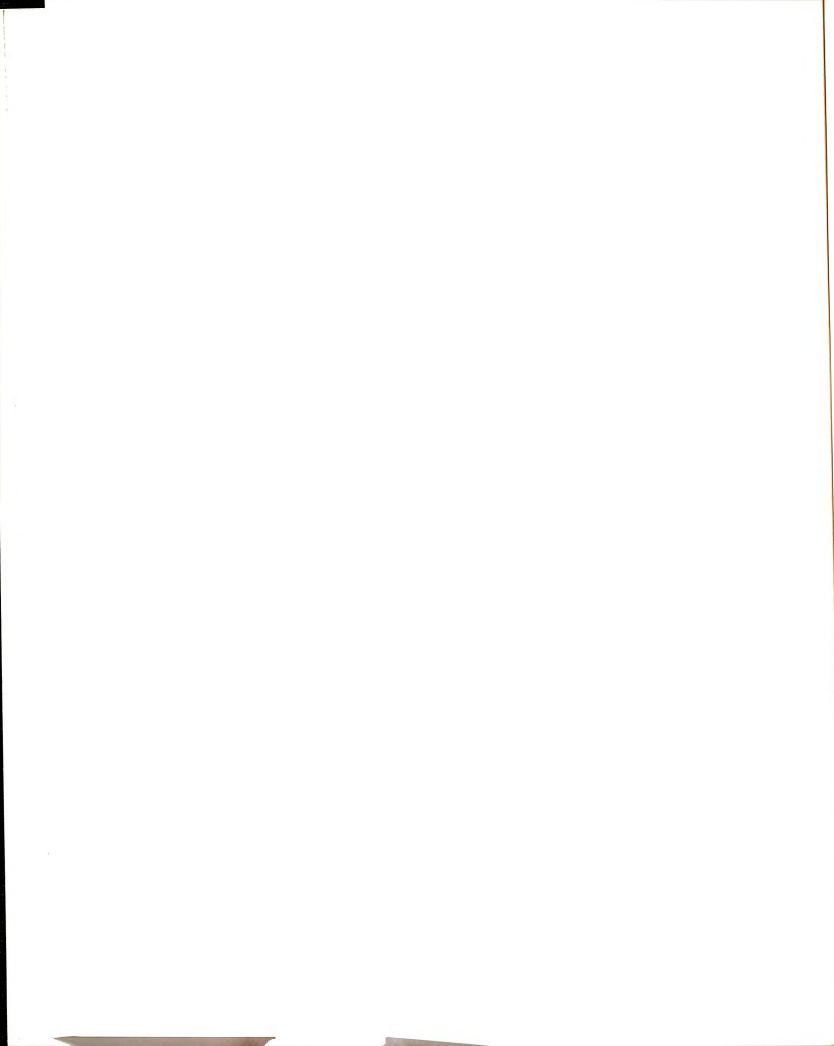
A subject might change his answer and / or method of approaching the problem at any time in the course of the experimental period. In this study, the final answer and / or process that the subject agreed to be his best judgment was taken as the data input point, no matter how many times the subject had changed his mind during the experiment.

Materials

The questionnaire used in this study consisted of seven problems. The first problem was the cab problem researched by Kahneman and Tversky (1973). The second was the same problem except that the base rate was modified to become a causal one (Tversky and Kahneman 1980). The third and the fourth were the engineer-lawyer problem with the description about Jack and about Dick. The fifth problem was the same as the third but with an extreme base rate of 1:99. The sixth problem was a base rate problem about the performance of a machine. The last problem was a mathematical Bayes' problem stated in linguistic form. Unfortunately, due to some unnoticed typing errors, this question was discarded from the final analysis. A copy of the questionnaire can be found in the appendix.

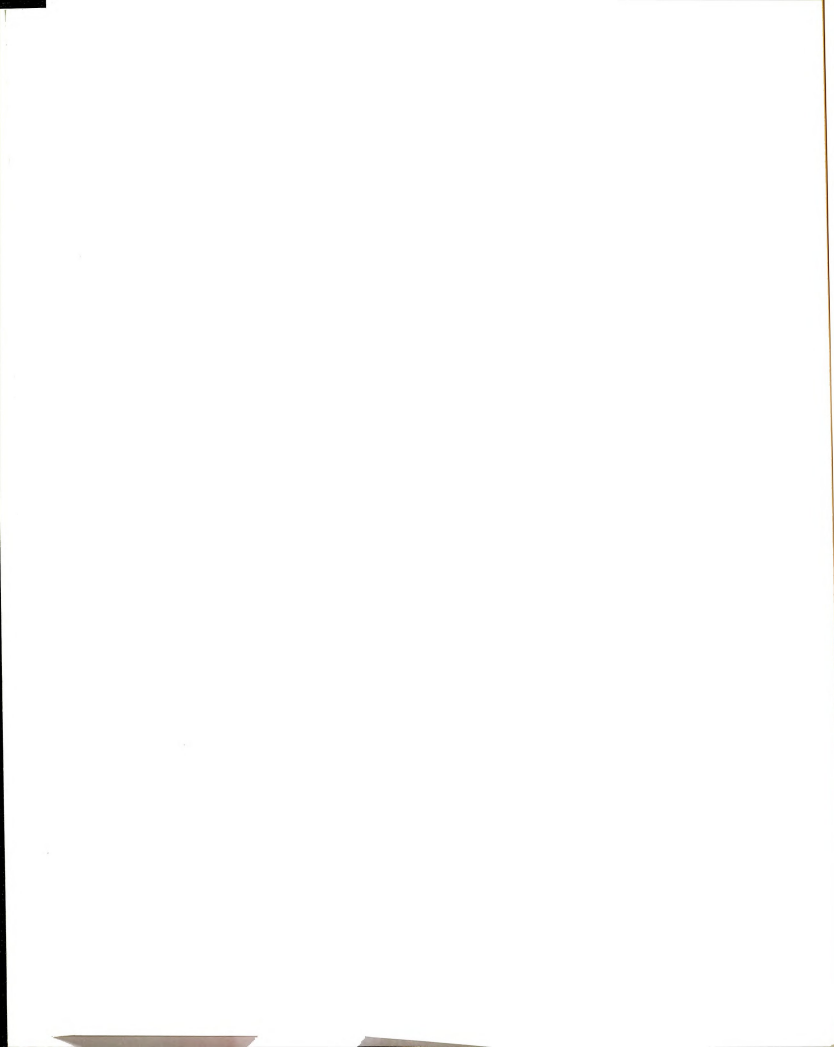
Specific Research Questions

The main purpose of this research is to restudy the base rate problem using the methods of thinking aloud as well as the research interview. It is hoped that these methods can reveal further the reasoning or thinking processes of the subjects in order to better understand how people judge.



We wanted to categorize people's thinking methods into several distinct and meaningful ways by studying the verbal protocols of the subjects. Therefore instead of looking at the responses from a perspective of checking whether or not subjects make use of base rate, we can actually see how subjects shift from one type of thinking category into others among different types of problem contexts. Using these descriptive categories, we can attempt to answer statistically (using Chi square) the following specific research questions:

1. Does the causality of the base rate affect people's mode of judgment? (by comparing problems A and B)
2. Does the diagnosticity of information affect people's mode of judgment? (by comparing problems C and D)
3. Do people judge differently to a base rate with high extremity (i.e. 99:1)? (by comparing problems C and E)
4. Do people judge differently between problems of a social context and that of a physical context? (by comparing problem F to A or B)
5. Do males judge differently from the females?
6. Does age affects the way people make judgment?
7. Do students' major areas affect how they would judge?
8. Do people who have learned Bayes' rule judge different from those who have not?

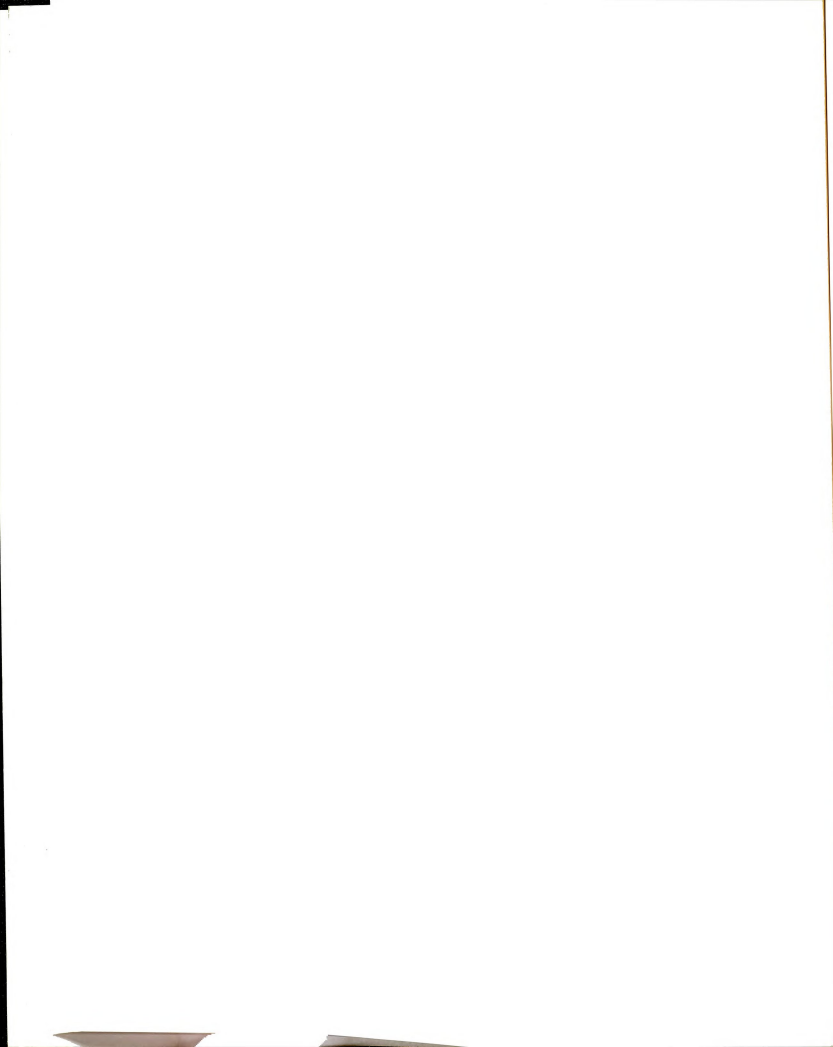


By comparing the numerical probability responses between our Chinese subjects and those of the American or Israelian subjects in past research, we can also get a rough idea about whether culture would make a difference in the response to base rate problems.

Analytic Method for the Verbal Protocols

Think-aloud data have been most useful in tracing the sequencing of information processing in problem solving. In our study, however, the critical information are very few, namely the base rate and the diagnostic information. It was found from our pre-analysis that the think-aloud data are not particularly illuminating in our case because the sequencing of information is not very important (e.g. as contrasted against playing chess or performing operations to control a factory boiler). Moreover, most of our subjects were not capable to verbalize very well about their thinking while solving the base rate problem. It was decided that our study would depend mainly on the verbal data from the research interview while those of the thinking aloud session only supplement concurrent information about what was happening while the subjects were solving the problems.

The analytic method for the verbal protocols in search of meaningful constructs is borrowed from what Marton (1981)

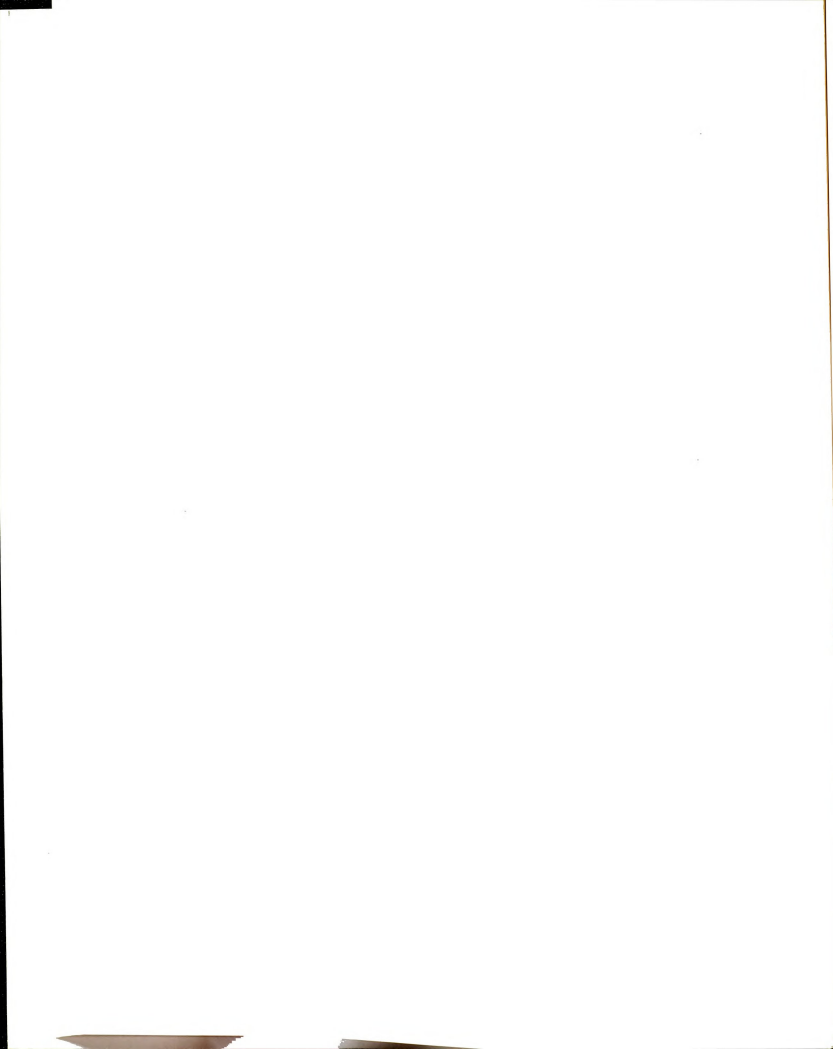


termed as "phenomenography". He explained it as "It is research which aims at description, analysis, and understanding of experiences; that is, research which is directed towards experiential description" (Marton, 1981, p.180). Such research is possible because the experience of reality have been repeatedly shown to be experienced in a limited number of qualitatively different ways. (See Gibbs, Morgan and Taylor; 1980 for an overview)

Marton also called this kind of research as second order research, as distinguished from the first order research which tries to describe various aspects of the world. He gave an example: the first order research is like asking the question "Why do some children succeed better than others in school?", while the second order research is like "What do people think about some children succeed better than others in school?".

Applied to our study, the first order perspective is like asking "Why do people make wrong probability judgment?". Our research perspective is similar to a second order one by asking "What are the grounds for your best-made judgment?".

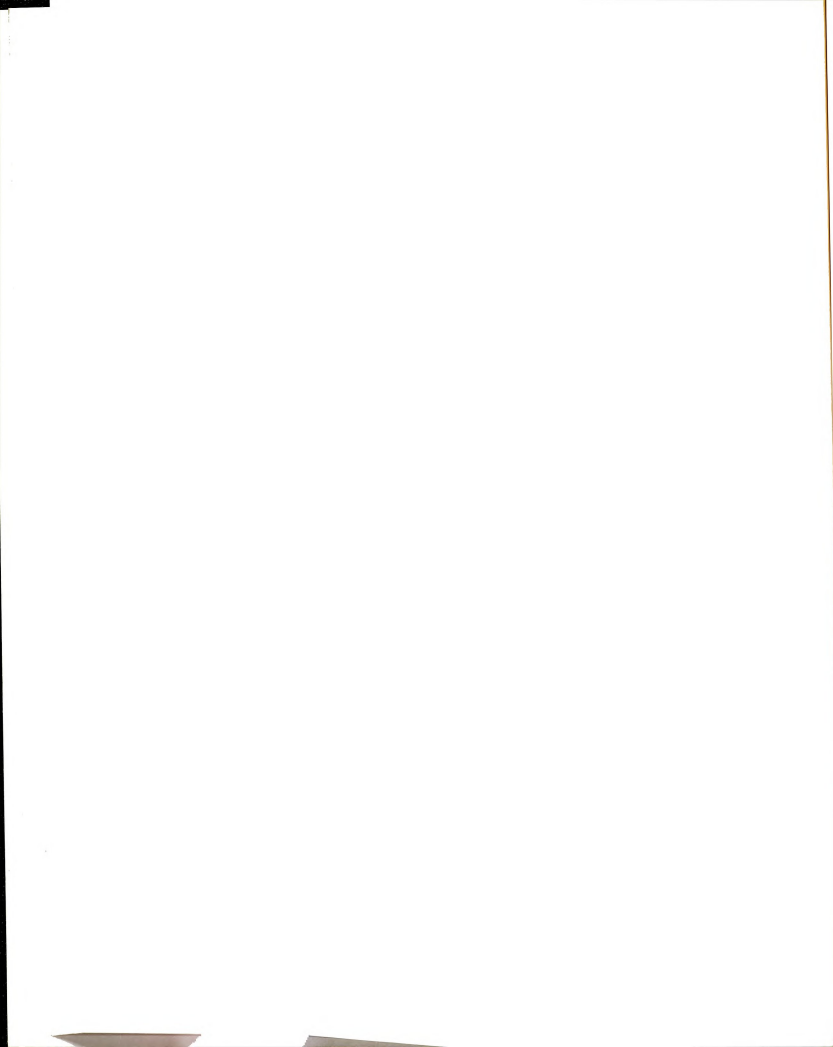
Two kinds of results are expected by this kind of research: the categories of description and the distribution of subjects over the categories. The categories



of description can be considered as abstract instrument for the analysis of concrete cases in the future. Or we can study a historical fact like individual X exhibit conception Y under circumstance Z.

In practice, the phenomenographic method applied in our case is as follows:-

1. The protocols of the individuals were first read, all comments relevant to enquiry are marked and identified.
2. The pools of comments thus obtained were then read for each problem across individuals.
3. Extracts were thus brought together into groups on the basis of similarity and the groups are delimited from each other on the basis of their differences. A higher order of meanings thus emerged are combined to form the categories of descriptions. The distinctive feature of this method is that "our analysis is dialectical in the sense that bringing the quotes together develops the meaning of the category, while at the same time the evolving meaning determined which of the categories are included or omitted" (Marton, 1984, p.55).



RESULTS

Our results are presented in several ways. First, the response of numerically judged probability is analyzed in the traditional way using the median as central tendency measure. By this we can compare our results with the other's results quantitatively. Second, the constructs or strategies established by qualitative methods are presented, and the rules used to form the constructs will be shown. Third, these strategies of judging will be used to make a statistical comparison between problems. The effect of problem contexts on the distribution of people's strategies of judging would be observed. Fourth, the effect of culture, sex, and major area on the preference of different modes of judgment will be presented.

A Traditional Analysis of Numerical Data

The numerical judgmental responses were submitted for a traditional quantitative analysis using median as central tendency measure. While this is not our main purpose, the results could be compared to other research based on these methods. The quality of data collection could be ascertained.

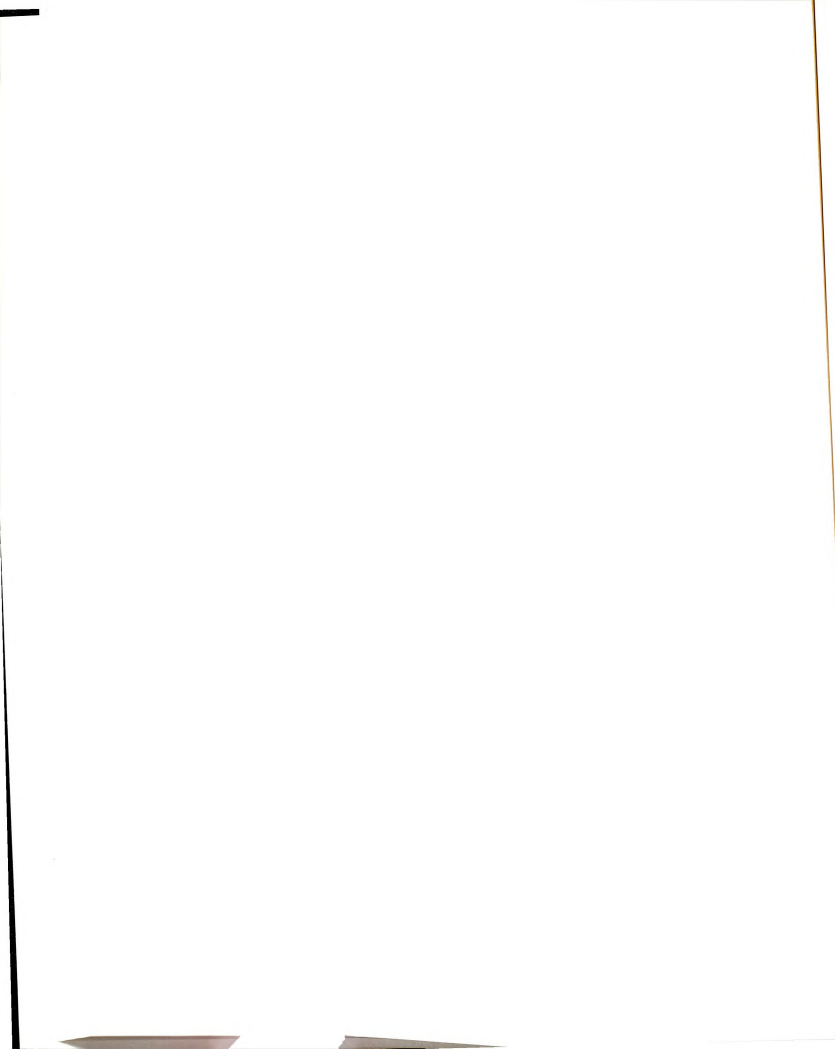


Table 1

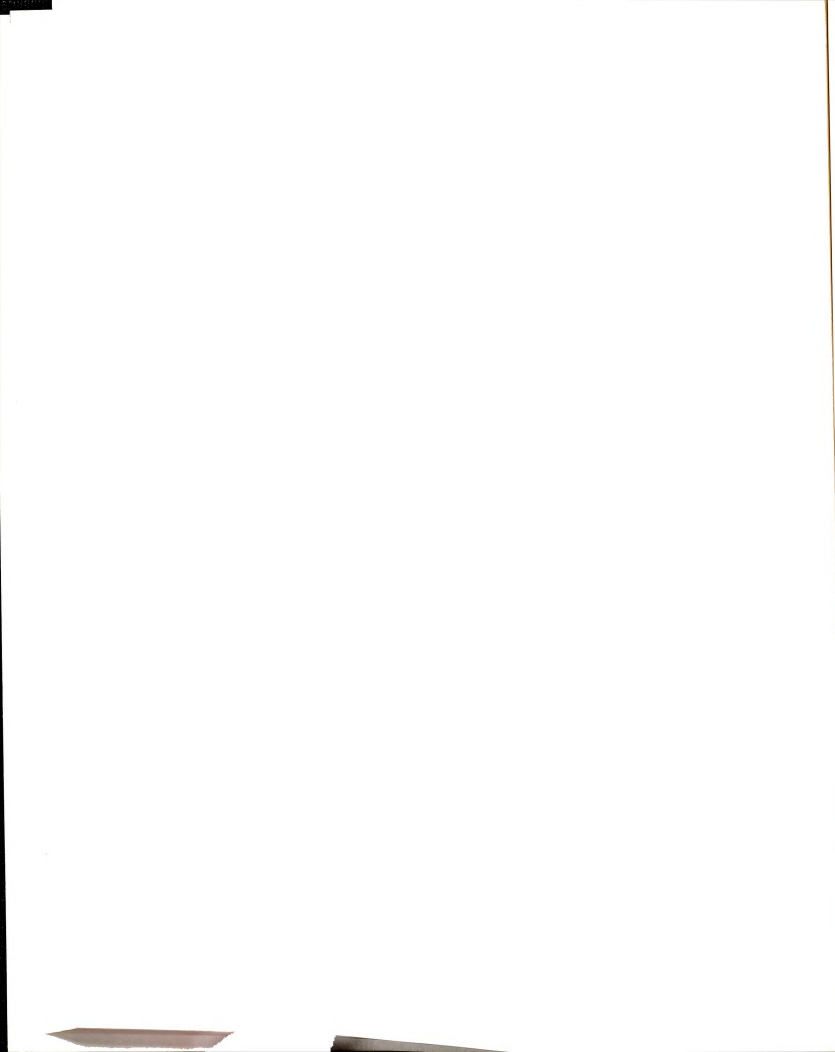
Quantitative Results based on Central Tendency

Mean, median and the corresponding Bayes' optimum for the judgment problems Studied

Problem	Mean	Median	Bayes' Optimum	Base Rate	N
A	55.7%	80%	41%	15%	39
B	54.2%	70%	41%	15%	39
C	65.9%	71%	n.a.	30%	40
D	32.4%	30%	n.a.	30%	38
E	39.6%	15%	n.a.	1%	40
F	34.5%	21%	21%	10%	39
G	38.6%	40%	14%	10%	14*

Note * : Valid subject size for problem G decreases due to defects in some questionnaires.

As we can see from Table 1, the median responses of the cab problems (A & B) are not close to the Bayes' estimate. The median for problem A and B is 80% and 70% respectively, while the corresponding Bayes' estimates are both 41%. There are no standard Bayes' estimates for the engineer-lawyer problems (C, D & E) because the diagnostic information for these problems are not written in explicit quantitative terms. The medians for problem F lies exactly on the corresponding Bayes' optimum, being 21%. For problem G, we have collected only 14 valid cases due to some typing errors



in the questionnaires for the earlier 25 subjects. In this problem, the median answer is 40%, quite apart from the 14% Bayesian optimum.

Results from problems A and B seem to repeat the findings of Tversky and Kahneman (1980). Our subjects' median answers, being 80% and 70%, are almost exactly the same as Tversky and Kahnemans' 80% and 60%. Since our subjects' answers fall short of the Bayes' optimum: 41%, they exhibit the usual fallacy. But the casual base rate in problem B seems to help shift the median response closer to the optimum. This shift is similar to Tversky and Kahnemans' result but a little less in magnitude.

Again, our results for the engineer-lawyer problems (C,D) are almost the same as those of Ginosar and Trope (1980, p.235). Our median answers are respectively, 71% and 30% while their results are 69% and 30%! This seems to confirm their findings that "base rates will be utilized to the extent that the usefulness of the individuating information for diagnosing category membership is diminished" (Ginosar & Trope, 1980, p.228).

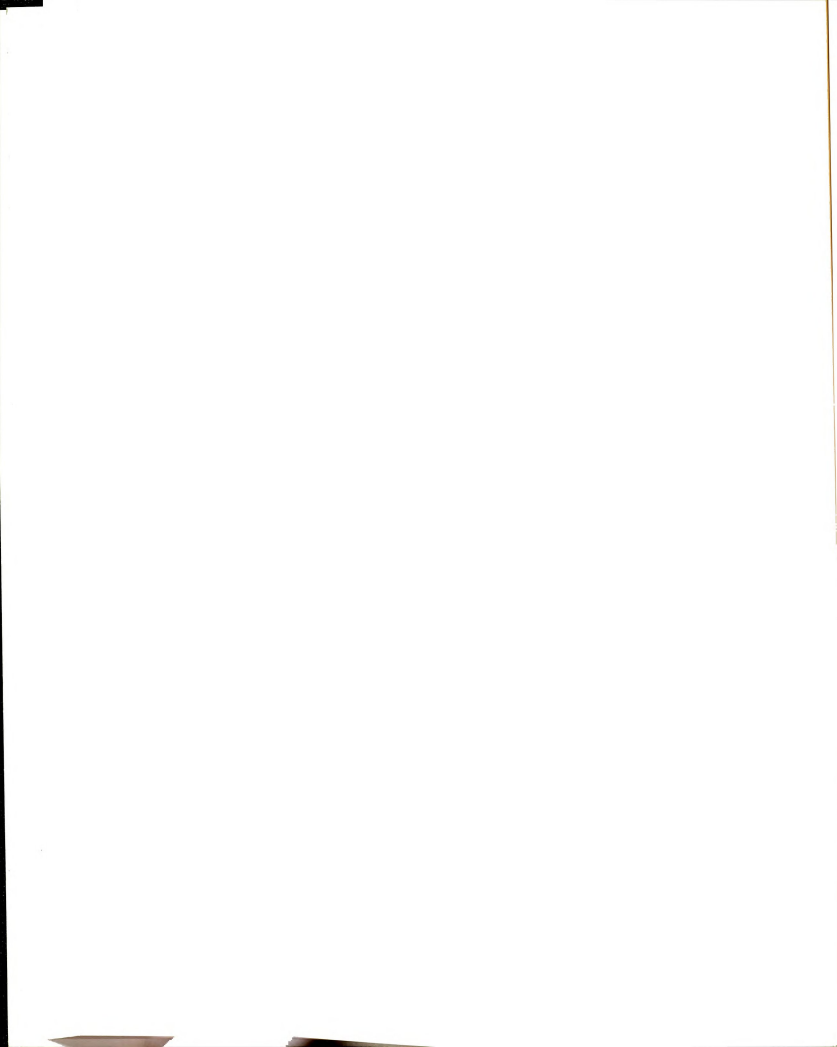
Problem E has a very extreme base rate (1:99), compared to problem C's (30:70). Our median response is 15%, diminished much from problem C's 80%. Because problem C and E are the same except for the base rate. The results seem to

show that under extreme base rate, people tend to use the base rate more.

Problem F is interesting because the median response of our subjects is exactly equal to the Bayes' optimum (i.e. 21%). The result seems to imply that when handling judgment of the pure physical realm, our subjects are on the average a perfect Bayesian! However, as we examine the distribution of responses carefully, (see Figure 3) most of the responses lie in the two modal regions of about 10% and 72%! (compare the base rate and diagnostic information :10% and 70%) Very few subjects' responses are close to 21%. The use of median response to summarize the results here is not justified.

Our last problem is a Bayes' problem in mathematical terms. Most subjects showed signs of difficulty in understanding or solving this problem. Unfortunately, due to a typing error in some questionnaires, valid sample size was reduced to only 14. It was decided to abandon this problem. Nevertheless, the median answer of 40% is quite far away from the optimal 14%

In sum, quantitative analysis using median seems to repeat major findings of other researchers. This give us some confidence that our data collection procedure is quite reliable. Results from problem F also highlight the fact



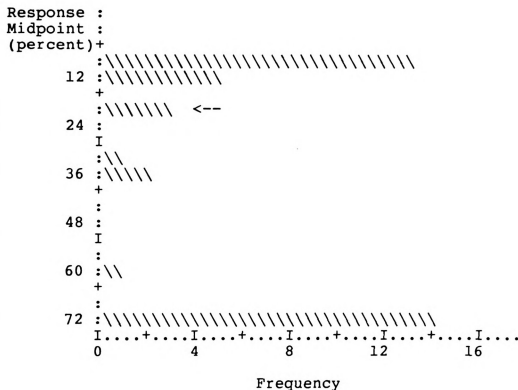
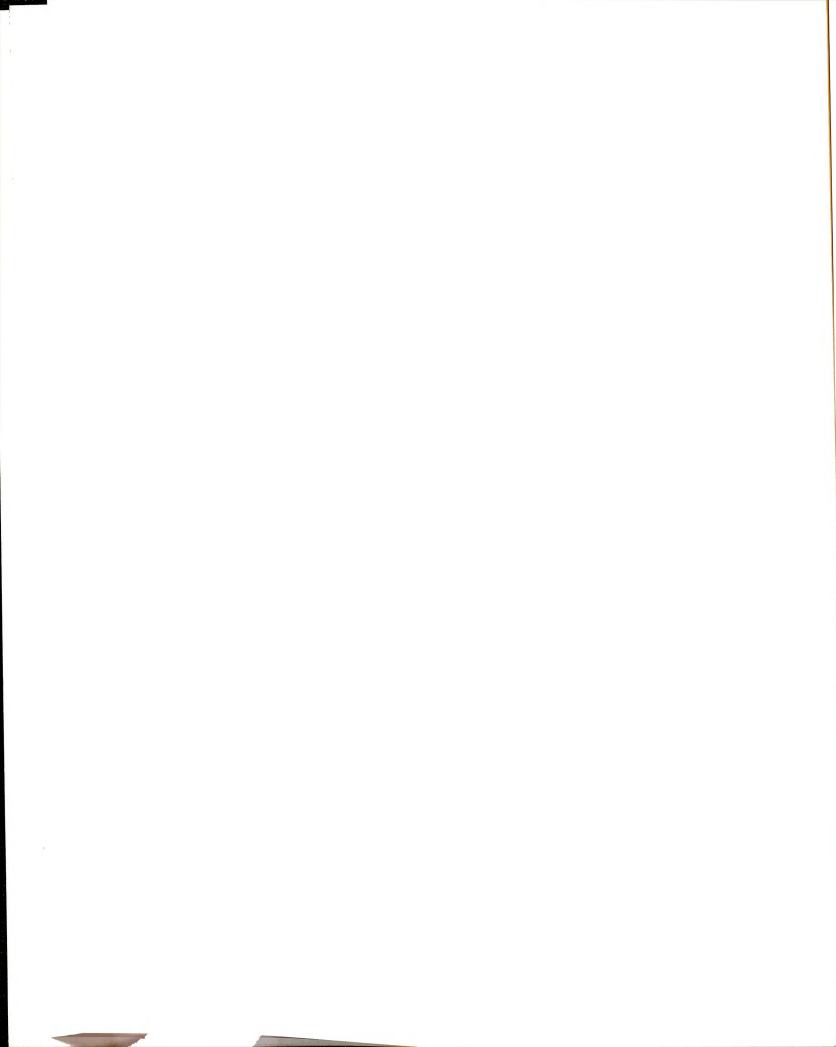


Figure 3: Histogram of the subjects' responses for problem F. Bayes' estimate denoted by the arrow sign <--.

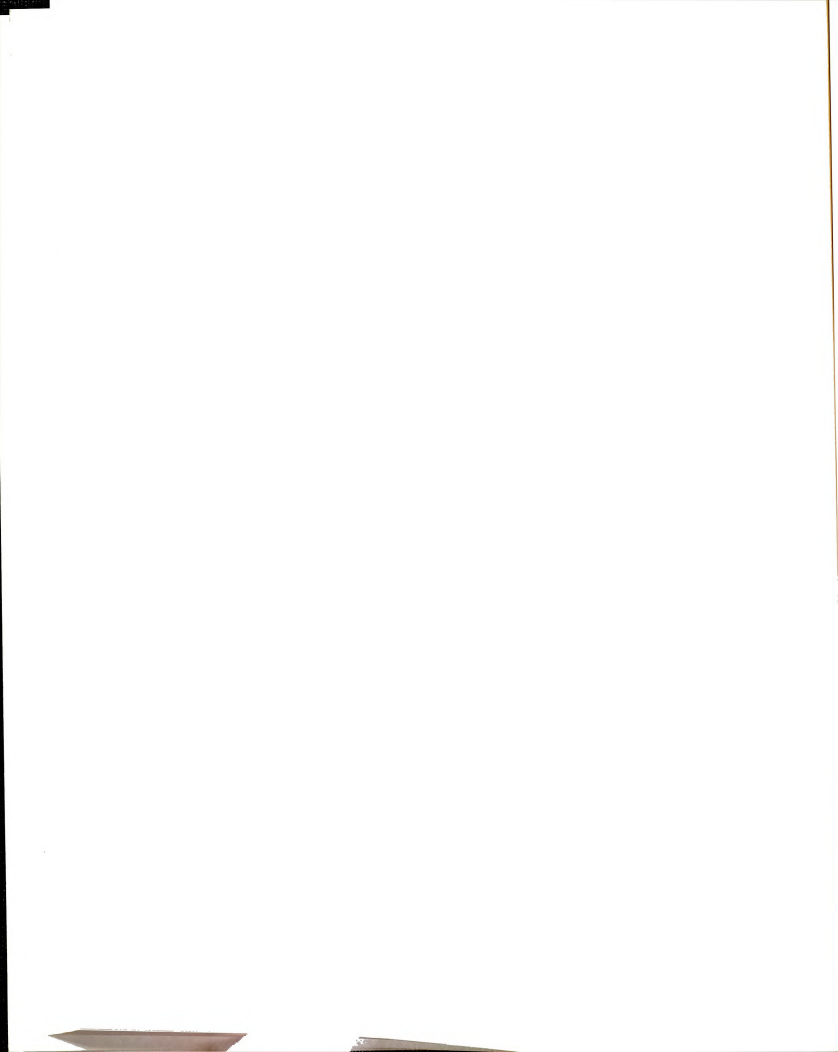
that using median response might not be justified as a summary of the responses. It also shows that using median response alone can easily neglect other distinct judgmental modes at work. For example, the distribution of responses in problem F was obviously bimodal.



Qualitative Categorization of Judgment Using Verbal Protocols

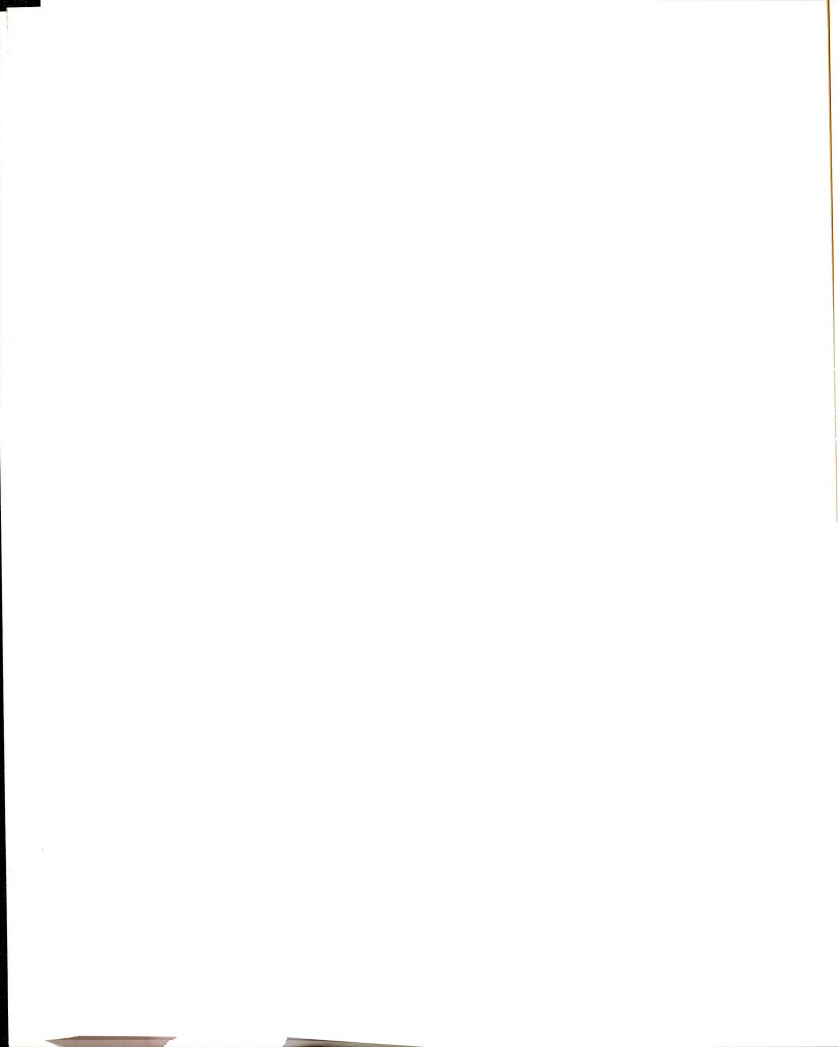
From our initial phenomenographic analysis of the verbal protocols from the interview, some qualitative distinct categories of judgment seems to stand out as probable research categories in our data. In the later phase of the data collection, fewer exploratory questions were asked and the data collection concentrated on finding out what particular way of judging the subject used. If the subject's judgment conformed with those earlier discovered categories, the amount of time asking related questions was shortened because the discovered categories act as a schema in understanding the current subject's way of judging. The shifting focus of data collection at various stages of research is called by Strauss (1987) as "theoretical sampling". Strauss explained this technique as one "whereby the analyst decides on analytic grounds what data to collect next and where to find them, ... so this process of data collection is controlled by the emerging theory ... When done well, this analytic operation pays very high dividends because it moves the theory along quickly and efficiently" (Strauss 1987, p.38-39).

Two core categories were finally decided for our set of data, they are the probabilistic mode and the intuitive



mode. In fact hardly any single pair of terms could fully represent the ways people judge. Other terms like scientific, axiomatic, mechanical or abstract may fulfill some of the descriptive functions not covered by the term probabilistic. For "intuitive" mode, words like lay, analytic, experimental or pragmatic might express some of the meanings not captured by the word "intuitive". Our terms are chosen because they seem to be more inclusive, less misleading and distinguishable from terminologies already in use by other researchers.

The term 'probabilistic' does not automatically means it is correct and intuitive does not mean incorrect. Probabilistic mode of judgment refer to the judgment which mainly utilizes the calculus of chance, e.g. the urn model, the axiomatic additive rule and multiplicative rule, or the conditional probability. Intuitive mode of judgment refer to the common everyday judgment which does not solely rely on axiomatic probability theories. No ready scheme or formula is used for judging. Instead, people using intuitive mode of judgment would consider the quality and relevance of the evidence, put higher weight on the particular case and on the present moment, might employ IF-THEN criterion or use narratives to fill in gaps of information. Descriptions about conceptual and the empirical indicators for the



probabilistic mode and intuitive mode of judgment are shown in Table 2 and Table 3 below.

Table 2

Indicators for Probabilistic Mode of Judgment

Concept indicators	Examples of empirical indicators
Subject to some calculus of chance	"It's a pure math or probability problem"
Use the "urn" model, judge solely by sample proportion	Use only the original ratio (i.e. base rate)
Employ some axiomatic probability theory e.g. the multiplicative rule, conditional probability	$0.9 \times 0.3 + 0.1 \times 0.7$ $80\% \times 15\%$ $0.1 \times 0.7 / (0.1 \times 0.7 + 0.9 \times 0.3)$

Using the above criteria for the two modes of judgment, we can treat the protocol of each problem answered by each individual as one data point and classify it as either belonging to probabilistic or intuitive mode. In a few cases where the verbal protocol was not very clear for categorization, the experimenter decides whether the judgment is predominantly probabilistic or intuitive. When all cases are settled, we can study quantitatively how the

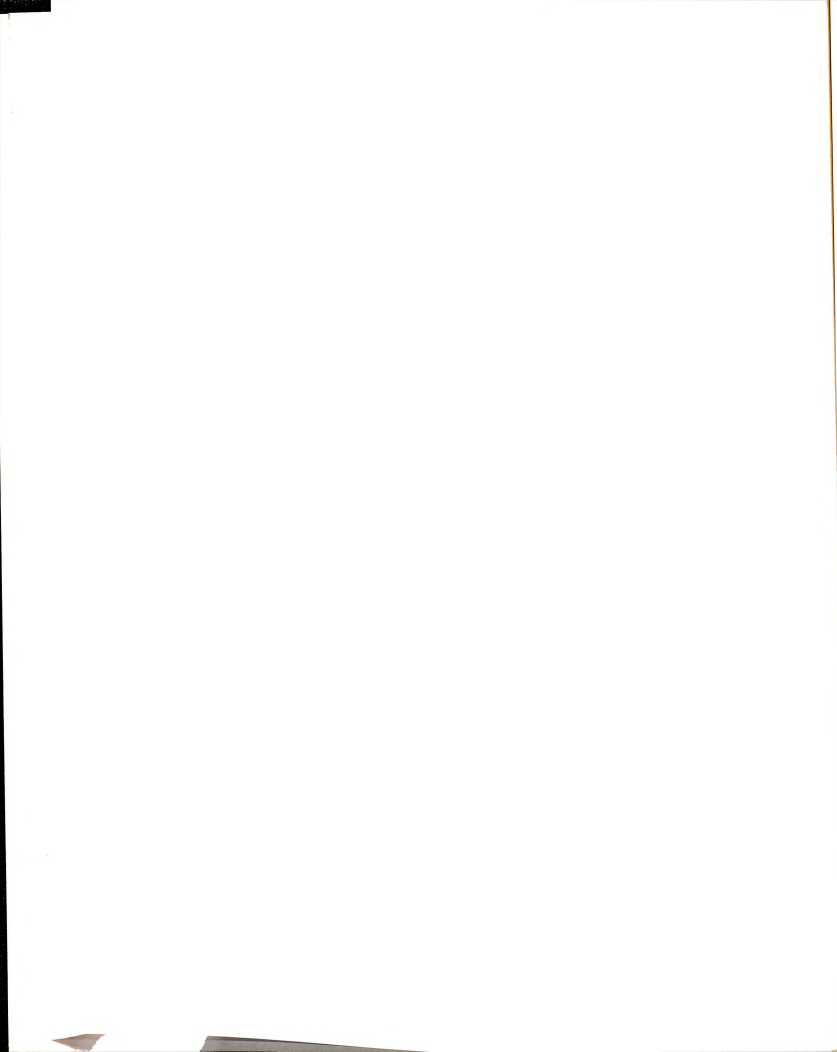


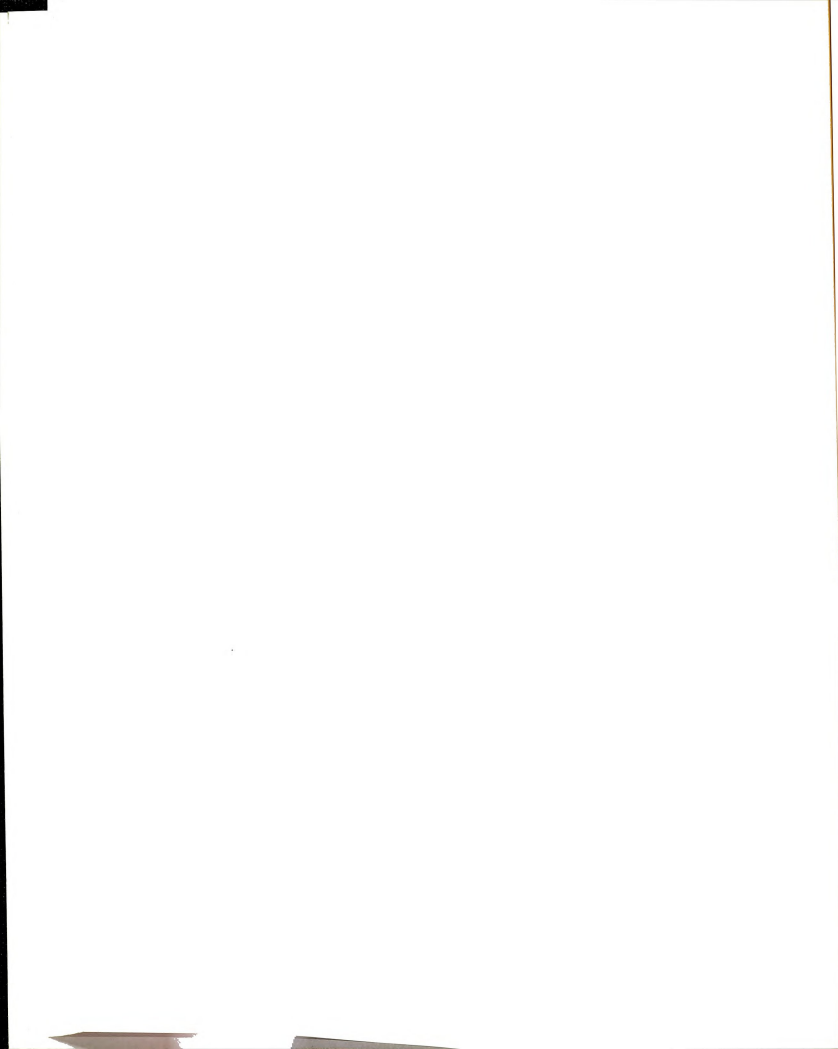
Table 3

Indicators for Intuitive Mode of Judgement

Concept indicators	Examples of empirical indicators
Does not rely on abstract axiomatic theory	"The description does not give us any information"
Consider the relevance and quality of evidence	"The number of cabs in city is irrelevant"
Generally put higher weight on evidence which is present or is about the individual case concerned	"The witness is more important, statistical data are not" "Past and present, no necessary relationship"
Employ IF-THEN criterion	"If no interest in social issues, hardly a lawyer"
Use narratives to fill in gaps of information	"Lawyer works in group, engineer works alone"

distribution of the modes of judgment varies as a function of problem context. Two examples showing the verbal protocols and the reasons for classifying them can be found in the appendix.

Reliability of coding. An independent judge was called upon to code the interview protocols according to the criteria from Tables 4 and 5. The resulted coding was checked against the original one coded by the experimenter.

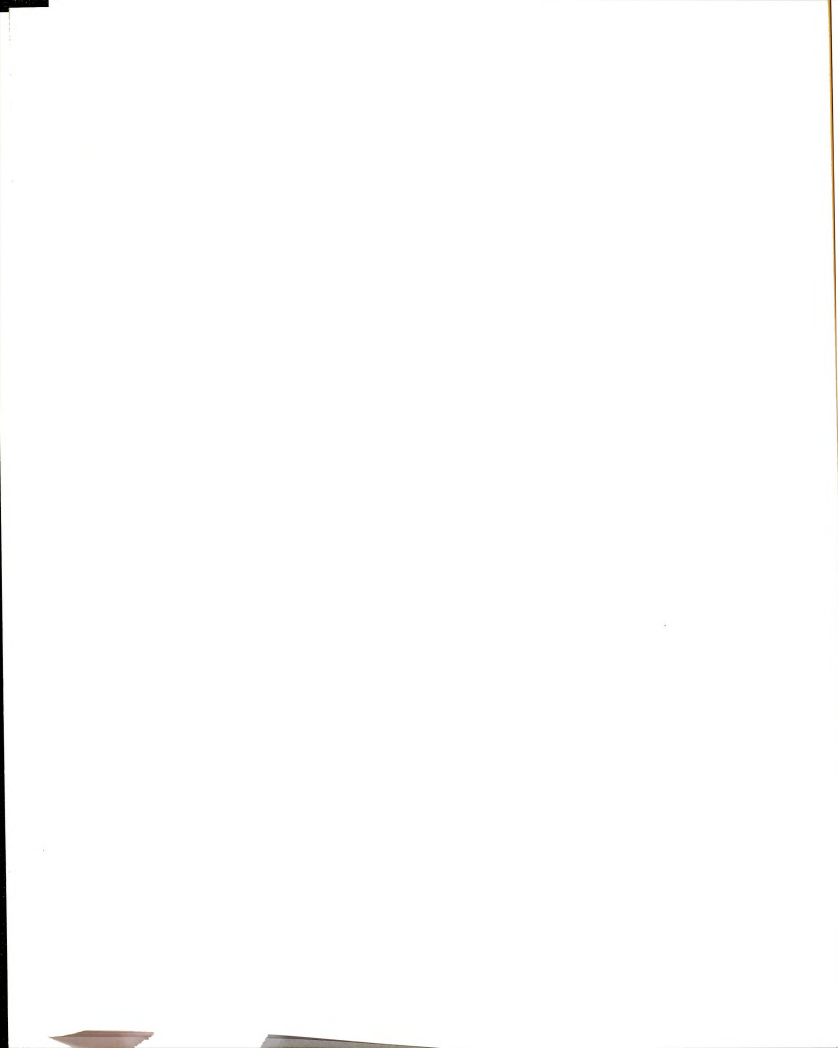


Extremely high reliability was observed. For problems A, B and F, only 1 out of 40 was missed. Two protocols were coded differently for problem C. Interrater correlation coefficients for problems A, B, F and C were respectively, 0.945, 0.947, 0.947 and 0.892. Exact match in coding was found in problem D and E. Accordingly, the coding criteria for the two categories of judgment should be very reliable.

Upon review of the discrepancies between the independent judge's coding against the experimenter's, the experimenter decided that the judge's coding was better and the data for final analysis was changed thereof.

Problem Contexts and Mode of Judgment: A Quantitative Analysis

If we compare the distribution of the number of subjects exhibiting the two modes of judgment between problems, we might be able to test how the problem type, or context would affect the distribution of modes of judgment. This comparison is similar to a pre- and post-test design. As outlined in our paragraph on research questions, we would like to test whether "causal" base rate, low diagnostic information, extreme base rate, and problem context (social vs. physical) affect the resulting distribution in modes of judgment. Obviously, since all we have are frequency data,



contingency table would be used. The McNemar test may best be used to test the null hypothesis that there had not been a change in the proportion of all subjects who used intuitive judgment (or equivalently, probabilistic judgment) (Conover, 1980, p.132). From this test we can also know whether intuitive or probabilistic mode of judgment have a higher proportion of changers between problems.

Before we begin to use the categories of judgment to analyze conditions related to the shift of judgment, we should examine whether the categories are relatively stable among problems. As shown in the table below, the phi coefficient indicates the correlation of the mode of judgments between problems. Higher correlation indicates greater stability. Among the contrasts of theoretical interests, the phi coefficients ranged from 0.282 to 0.837 and are all statistically significant. The mean phi was 0.550. Thus the average high phi coefficients indicates that the qualitatively coded modes of judgment are stable enough as a construct.

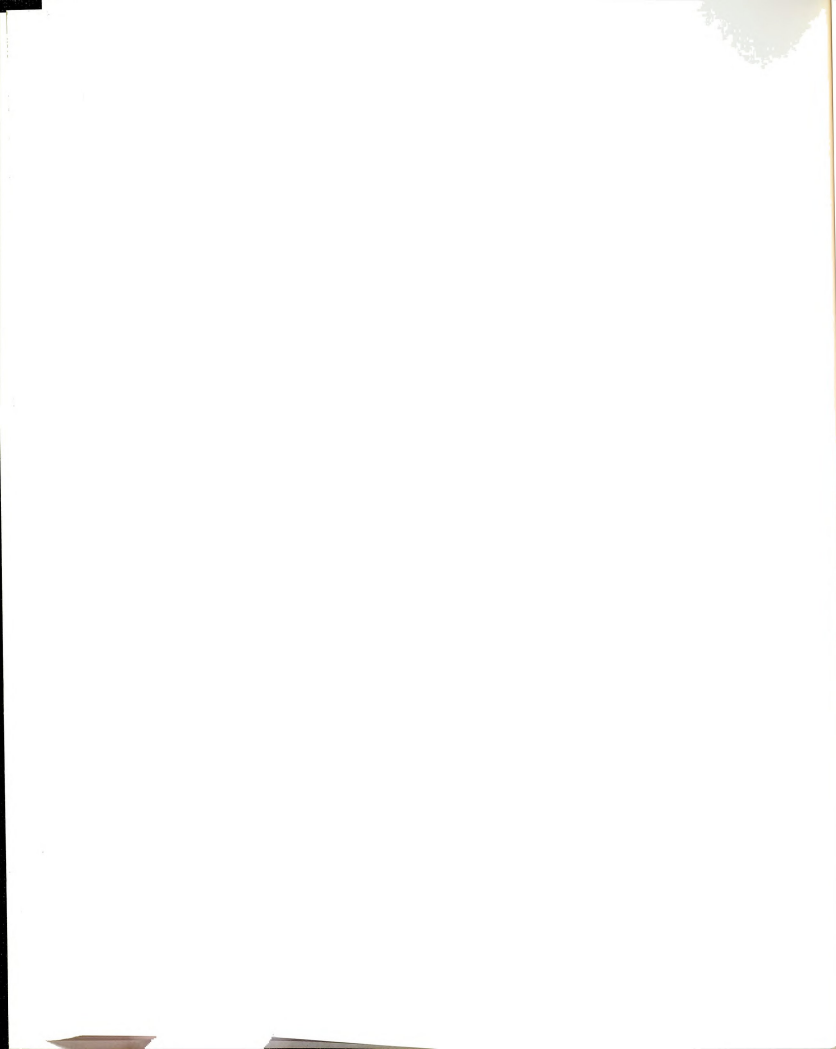
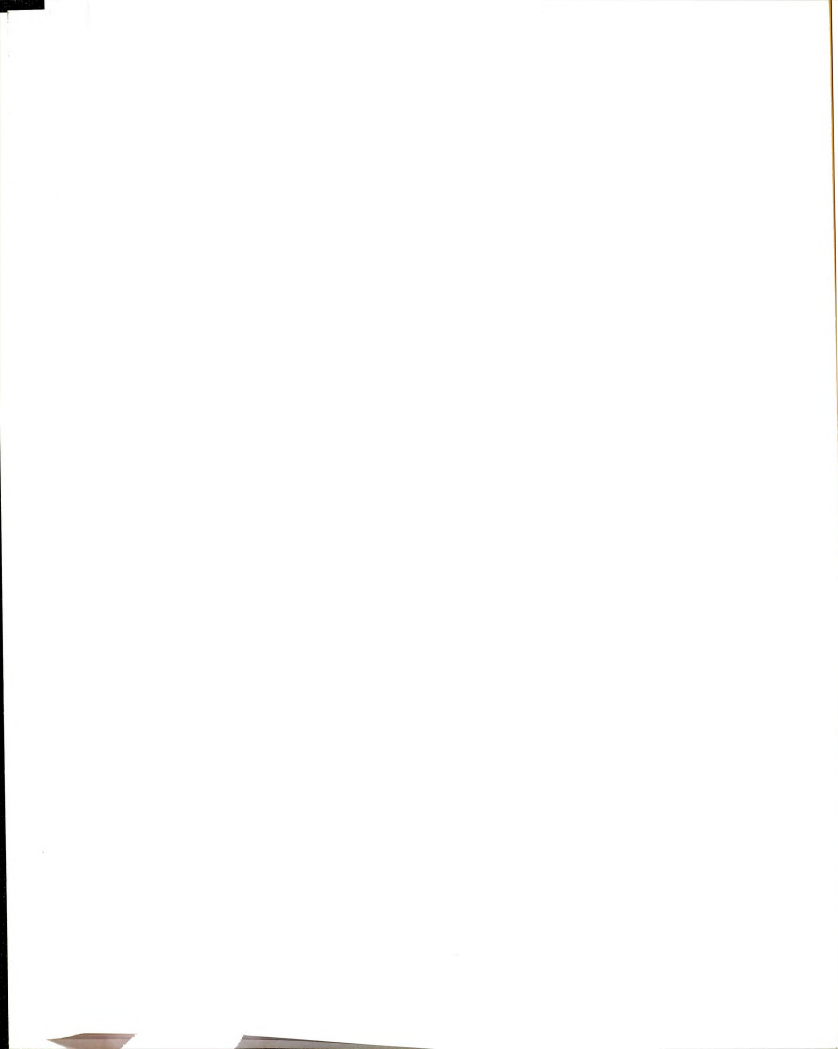


Table 4

Stability and Change under Various Problem Conditions

Stability and change were tested by the phi coefficient and the McNemar's chi square respectively.

		(B)		Phi Coeff.	McNemar's Chi Square
		Intuit.	Prob.		
	Intuit.	23	2	0.837	0.33
(A)	Prob.	1	13	($p < .001$)	($p > .05$)
		(D)		Phi Coeff.	McNemar's Chi Square
		Intuit.	Prob.		
	Intuit.	17	11	0.551	11.0
(C)	Prob.	0	11	($p < .001$)	($p < .001$)
		(E)		Phi Coeff.	McNemar's Chi Square
		Intuit.	Prob.		
	Intuit.	22	6	0.724	6.0
(C)	Prob.	0	12	($p < .001$)	($p < .02$)
		(F)		Phi Coeff.	McNemar's Chi Square
		Intuit.	Prob.		
	Intuit.	12	13	0.357	8.07
(A)	Prob.	2	11	($p = .014$)	($p < .01$)
		(F)		Phi Coeff.	McNemar's Chi Square
		Intuit.	Prob.		
	Intuit.	11	13	0.282	6.25
B)	Prob.	3	11	($p = .043$)	($p < .02$)

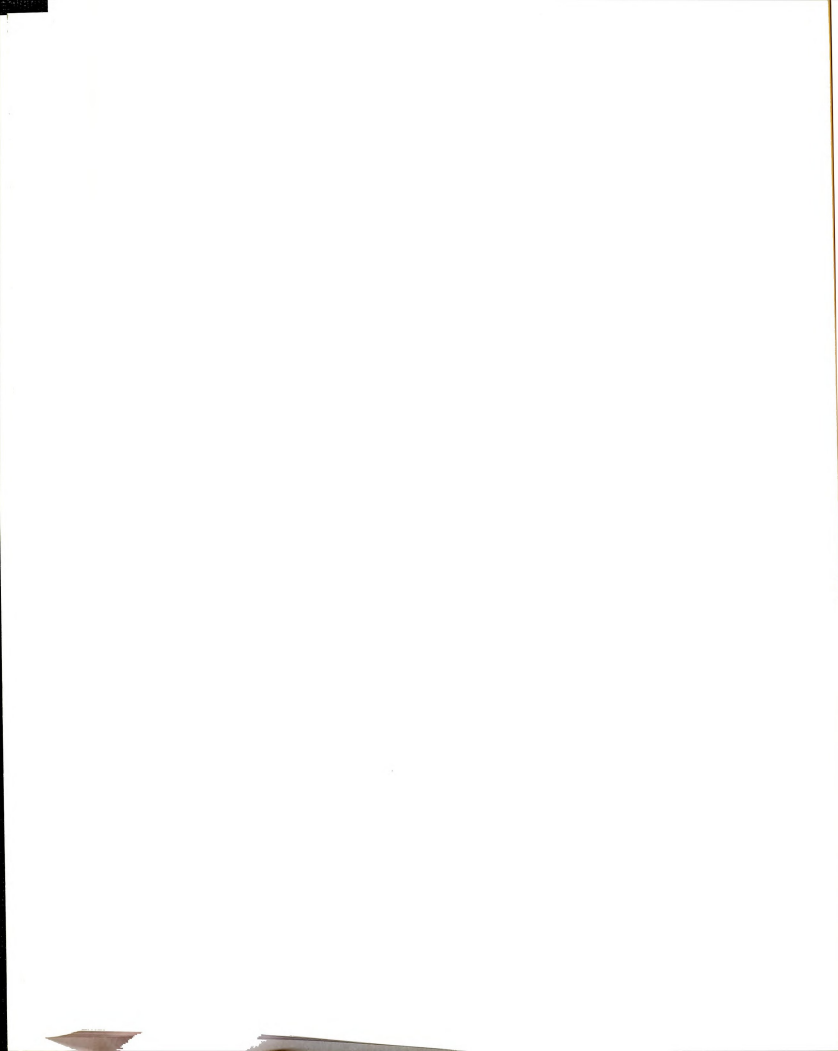


Causality and mode of judgment. Causality is supposed to be what makes problem B differs from problem A. Because in problem B, the base rate is given as accident rates instead of the relative cab size of the two cab companies. Past research showed that there was a drop in the median response from about 80% to 60%, closer to the Bayesian optimum 41%, when the base rate in the cab problem was given in accident rates (Tversky and Kahneman 1980). The drop was attributed to the causality of the base rate that "readily elicits the inference that the drivers of the Green cabs are more reckless and/or less competent than the drivers of the Blue cabs" (Kahneman et al. 1982, p.157).

In fact only 12 (30.8%) out of 39 subjects in our study gave the "causal" problem a different numerical answer than that of problem A. For the majority of subjects (69.2%), the "causal" base rate did not affect the way they made judgment.

With regard to the change in mode of judgment, only 3 subjects shifted their modes between problem A and Problem B. The change was not statistically significant (McNemar's $\chi^2 = 0.33$, $p > 0.05$).

The reasons given by the intuitive subjects who did not shift their judgment are summarized below. The subject S2B for problem B, for example, are denoted by S2B. The



case where there are over five subjects is not listed by individual subjects. Only the subject size will be given.

1. Past and present are independent, they have no (necessary) causal effect. (S2B,S9B,S40B,S24B,s29B)
2. Witness is more important, more reasonable. (Reasoning similar to that for problem A). (N=12)

Five subjects who remained intuitive but nevertheless affected by the causal base rate lowered their numerical estimates. They reasoned as:

1. Accident rate has some influence. Lower the witness's reliability. (S33B,S26B,S1B,S28B,S38B)

For all subjects who shifted their judgment, their reasonings were:

1. We should consider the accident rate. I used mathematical calculation. (S27B)
 2. This concerned the occurrence of accidents, we have to consider this information. (S25B)
- Accident rate is just background information, witness is more important. (S10B)

Subjects using probabilistic mode of reasoning usually have no specific reasonings for using the same strategy. They considered the mathematical calculation as the appropriate method to answer these problems.

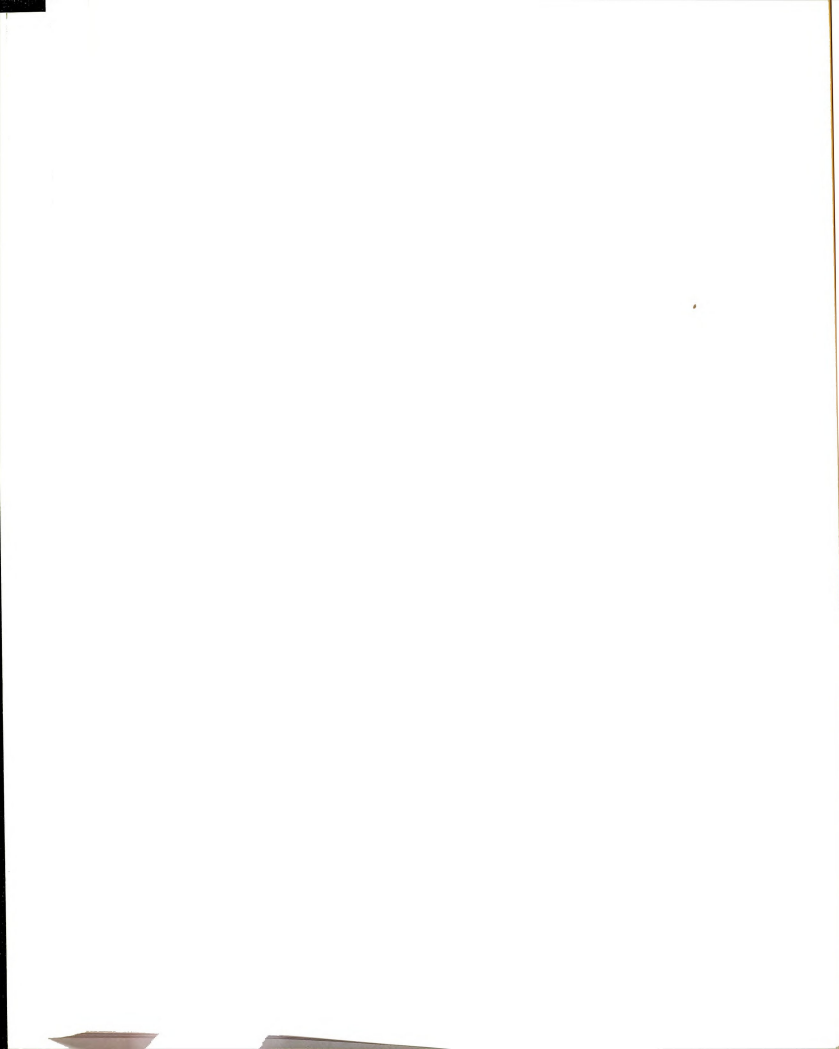


Table 5

Effect of Causal Base Rate on Judgment

		(B)	
		Intuitive	Probabilistic
(A)	Intuitive	23	2
	Probabilistic	1	13

Diagnosticity and mode of judgment.

The diagnostic information in problem C and D differ in their diagnosticity. The description of Dan was intended to convey little diagnostic information. Past research (Ginosar & Trope 1980) showed that the engineer-lawyer problem's median response dropped from about 70% to 30% when the low diagnostic description of Dan replaced the high diagnostic personality description of Jack.

Our study showed that 16 subjects (40%) resorted to the base rate of engineer (i.e. 30%) to answer problem D, because they thought the description was vague or contained little information. Based on the same reason, 5 subjects (5%) gave an either-or (i.e. 50%) as the answer. Together (N=21) of subjects described Dan's description as

providing no information to decide for a career between engineer and lawyer. Interesting enough, 6 subjects (15%) thought that Dan's description now seemed more like a lawyer, the percentage they gave ranged from 5% to 20%. These subjects explained the choice by saying since Dan is of high ability, high motivation, it sounds like an achieving young lawyer. In addition, two of them believed an engineer works alone, so the description of "He is well liked by his colleagues" doesn't fit for an engineer. With regard to the shift of judgmental mode, it was found that out of 28 subjects using intuitive mode for problem C, 11 (39.3%) of them shifted to probabilistic mode. They choose the sample ratio (30%) as answer, noting that the description was too vague. For those using probabilistic judgment in problem C, all of them remained as probabilistic for problem D. Eight (72.7%) of these subjects admitted that the vague description confirmed their use of the sample ratio as answers. The shift of mode here was statistically significant, McNemar's chi-square= 11.0, $p < .001$.

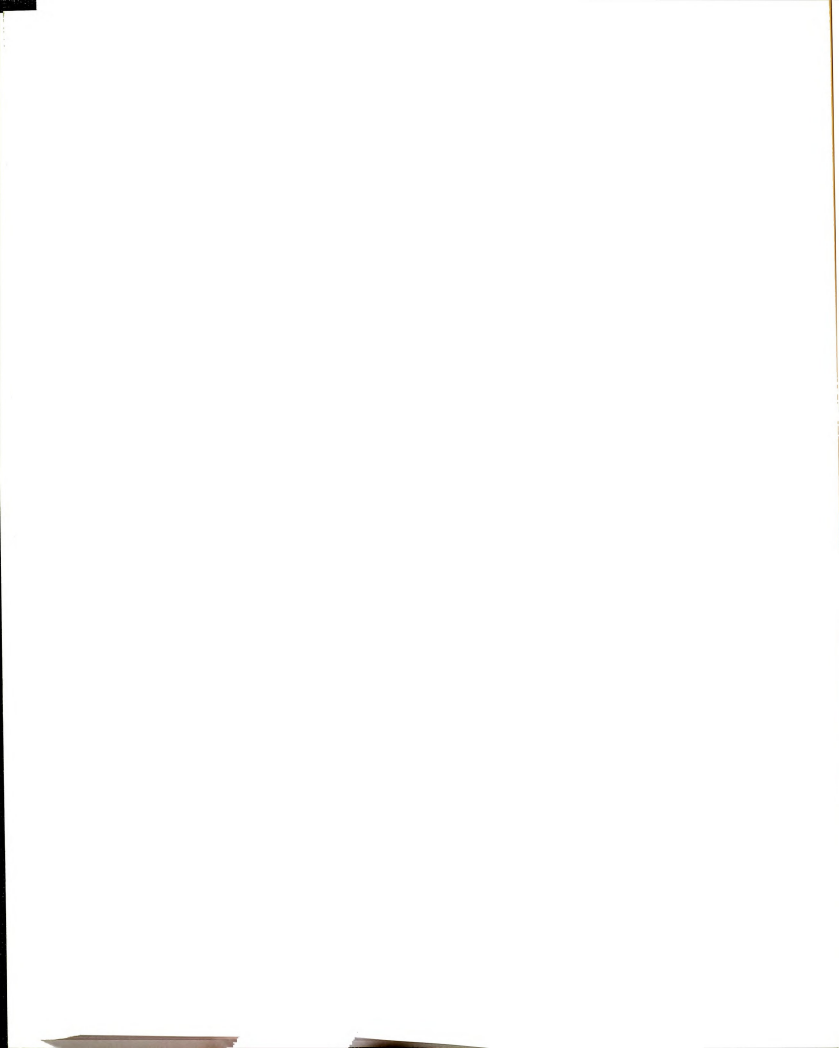


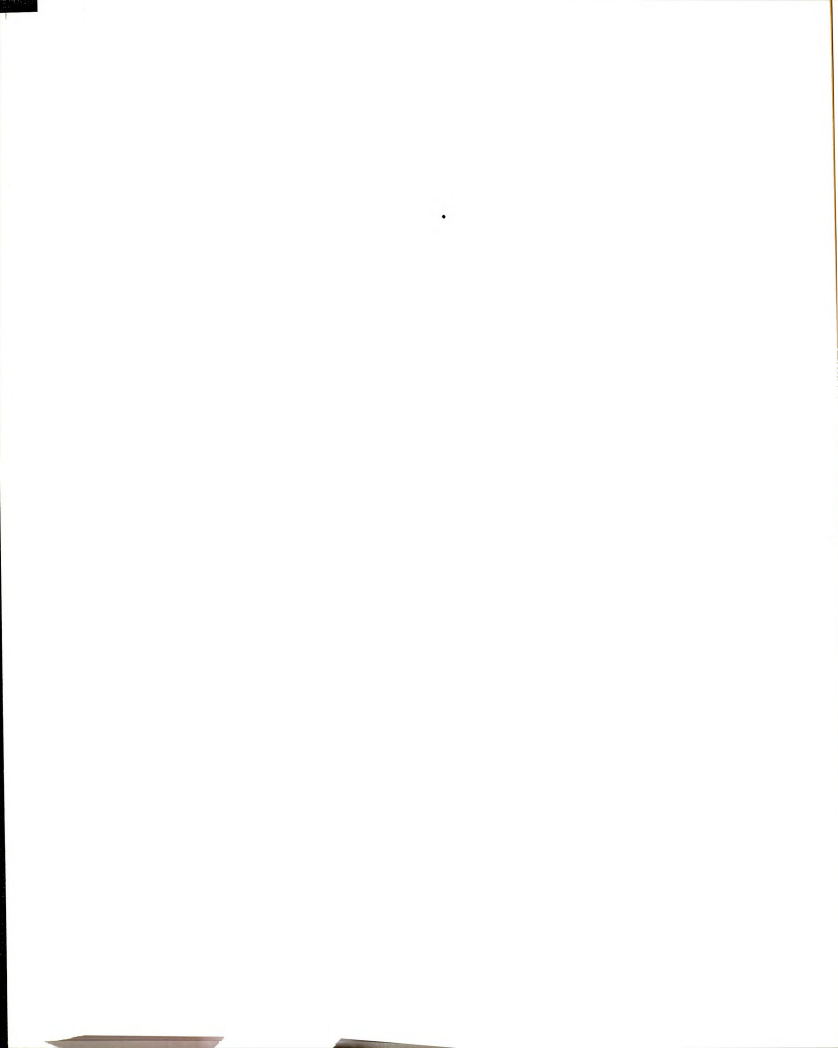
Table 6

Effect of Diagnosticity on Judgment

(D)

		Intuitive	Probabilistic
(C)	Intuitive	17	11
	Probabilistic	0	11

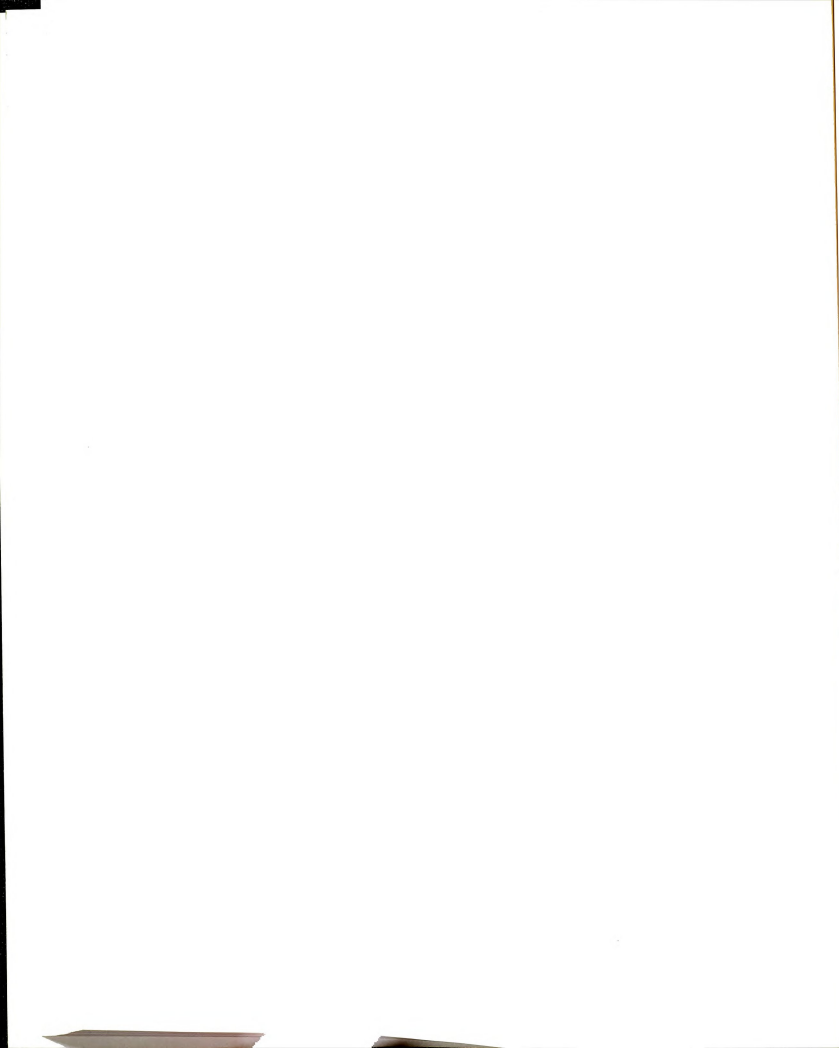
Extreme base rate and mode of judgment. The results of adopting extreme base rate (e.g. 90%/10% or more) have produced equivocal results (see the review by Borgida and Breke, 1981). Our study have found that 19 (67.9%) of the 28 intuitive judge in problem C were affected by the extreme base rate of 99%:1% in problem E. 7 (25.0%) of those intuitive judges decreased their confidence about the description as belonged to an engineer, although the probability they gave was still over or equal to 50%, indicating also that the description is still an engineer. They claimed that extreme base rate had an effect, but the description was still like an engineer. Five (17.9%) subjects used the 1% as ground level but gave an answer slightly higher than 1% (i.e. values ranged from 5 to 20%)



to indicate their belief that the description looked like an engineer, but the probability should not be as low as 1%. Another 7 intuitive subjects (24.1%) shifted the mode of judgment to probabilistic, 6 of them adopted the sample ratio (1%) as answer while the remaining one made some calculation with the use of probability theory. 9 intuitive subjects (33.3%) in problem C gave the same answer to problem E, unaffected by the extreme base rate. Their main reasonings are listed below:

1. I focus on the character. (S7E)
2. Just like the taxi problem, ratio does not have much meaning. (S40E)
3. Aged 45. a lawyer with no interest in social and political affairs, not likely. (S29E)
4. It has relation with the character, not number. (S22E)
5. I base my judgment on the description, statistical data has no relation. (S31E)
6. The information is so strong. (S6E)
7. The description can hardly be a successful lawyer. (S9E)
8. No social and political interest, like math puzzle, it is an engineer. (S37E, S35E)

For the probabilistic judges of problem C, all of them remained probabilistic for problem E. The overall shift of



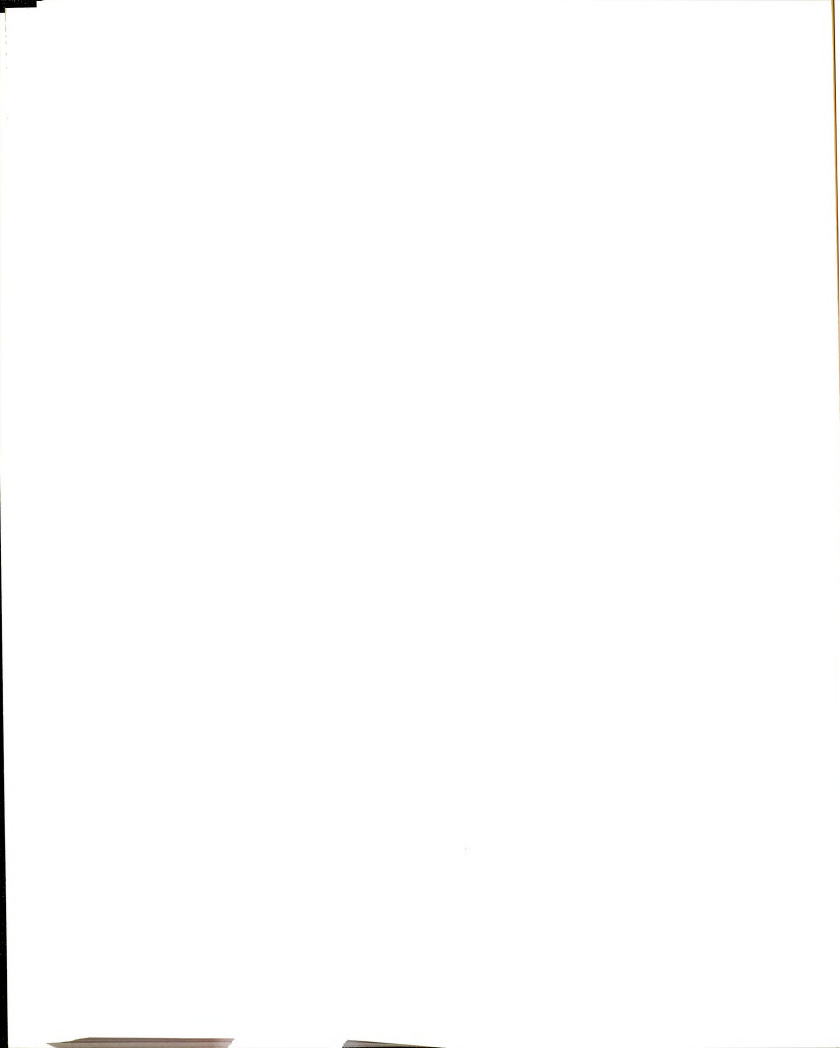
judgment between subjects in problem C and problem E is statistically significant, McNemar's chi-square=6.0, $p < .02$.

Table 7

Effect of Extreme Base Rate on Judgment

		(E)	
		Intuitive	Probabilistic
(C)	Intuitive	22	6
	Probabilistic	0	12

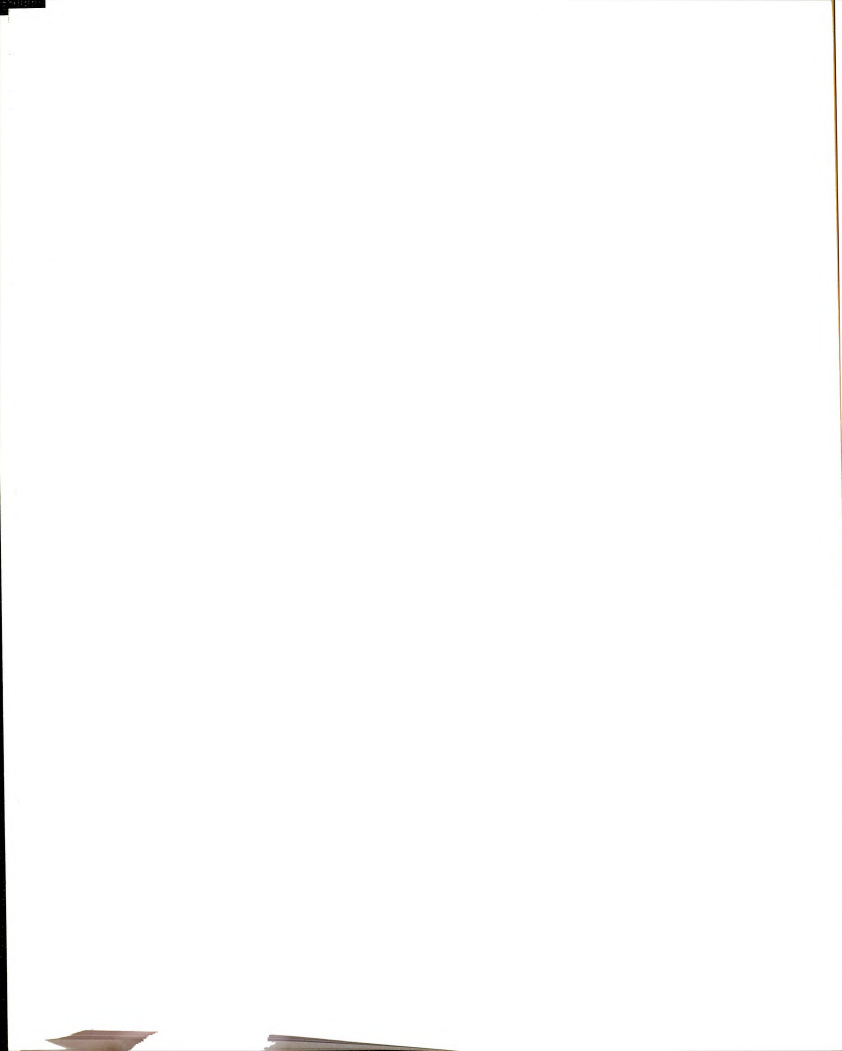
Problem context (social vs. physical) and mode of judgment. Base rate fallacy is often documented under circumstances of life-like problems, e.g. judging on witness' reliability in court and inferring a person's career from one's character. One might be curious how the subjects, usually college students, might err on the base rate problems and be able to study advanced mathematics courses on the other hand. A hypothesis might be for students with some backgrounds in introductory probability theory, they might be more prone to intuitive judgment for life-like problem, and to probabilistic judgment for more



mechanical, physical problem. Our sample of subjects are particularly appropriate for testing this hypothesis, as all of them have at least learned some elementary probability theory in high school. Maybe there is some significant difference between the social and physical world as perceived by people that they would use different mode of judgment for the two worlds.

Problem F is designed to be a physical, mechanical problem that is concerned with the accuracy of a machine with computer vision on a testing document which contains ellipses and circles of different proportions. This problem is highly comparable to problem A or B because the latter problems is concerned with the accuracy or reliability of the witness while the cabs of different colors have a different proportion or a different prior accident rates. The crucial difference between problem E and problem A or B is the problem context, for the former is life-like and in the social world; while the latter is mechanical and within the domain of the physical world.

Our results confirm our prediction, the shift of judgmental mode, mainly from intuitive to probabilistic, is statistically significant. McNemar's chi-square was 8.07, $p < .01$ between problem A and F and 6.25, $p < .02$ between problem B and F.



In problem F, many subjects who turned to probabilistic mode did not give specific reasonings for their doing. Nevertheless, the significance of the difference between a human affair and a machine in affecting judgment as perceived by our subjects can be traced by the following clues. The first five reasonings came from intuitive subjects.

1. It is a machine, more mechanical, therefore it's probability should be 70%. (S20F)
2. It is a machine, not a man. It does repeated actions, its error rate should be the same. (S22F)
3. Computer is rather 'dead' thing. When the computer has made an answer, the original document's ratio does not reflect the error. (S23F)
4. Because it is a machine, I would trust more about its reliability, (S27F)
5. It is mechanical, more mathematical, it is different from the previous personality problem (S39F)
6. This is a machine, not a human being. Therefore it is a pure math problem. (S9F, a probabilistic subject)

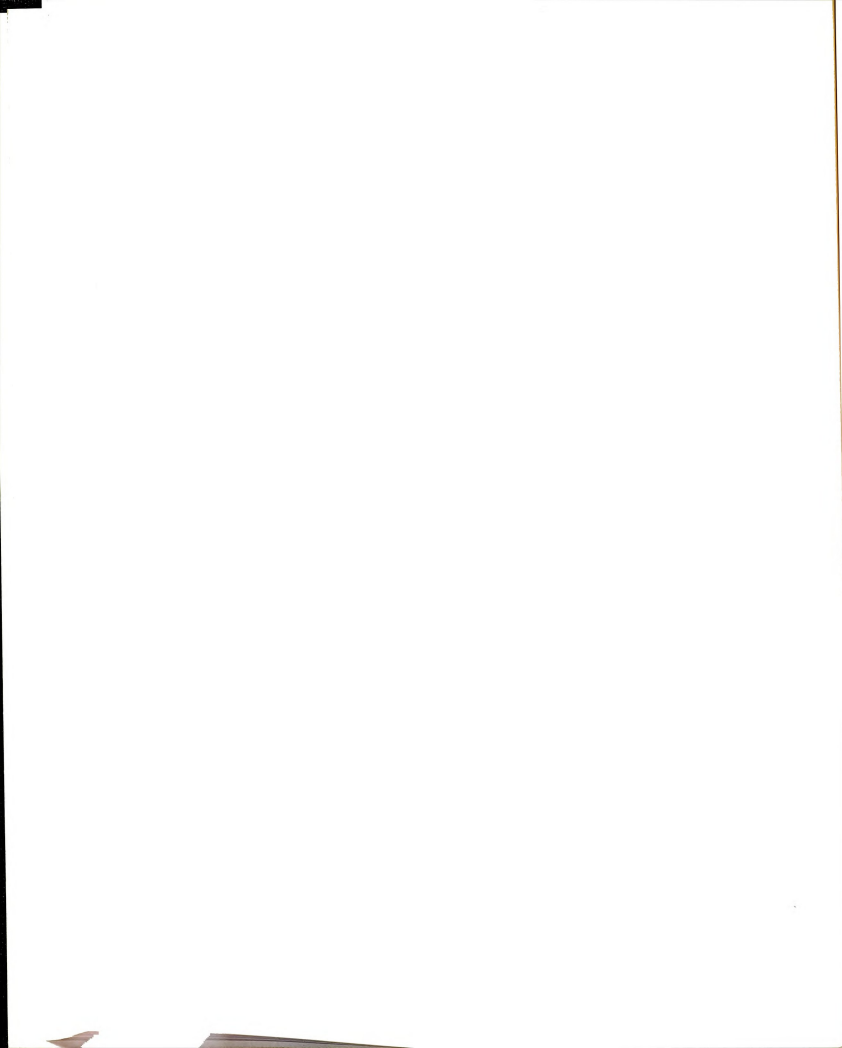


Table 8

Effect of Physical Context on Judgment (I)

(F)

		Intuitive	Probabilistic
(A)	Intuitive	12	13
	Probabilistic	2	11

Table 9

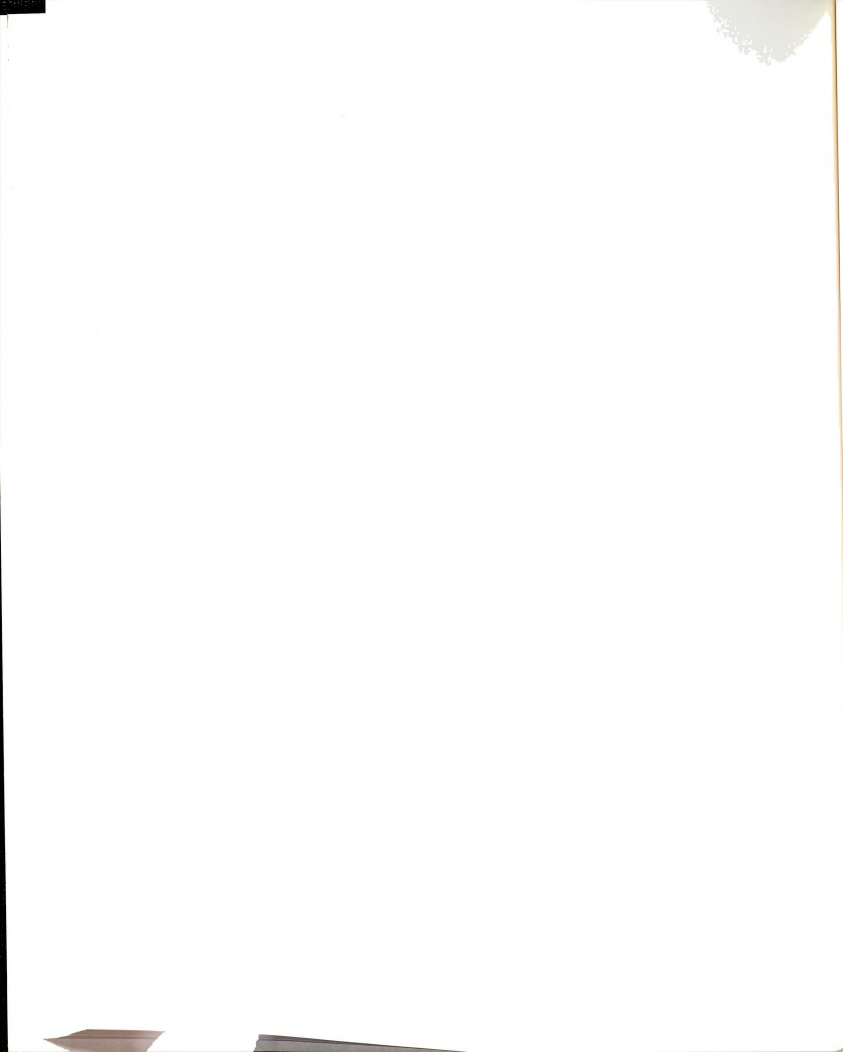
Effect of Physical Context on Judgment (II)

(F)

		Intuitive	Probabilistic
(B)	Intuitive	11	13
	Probabilistic	3	11

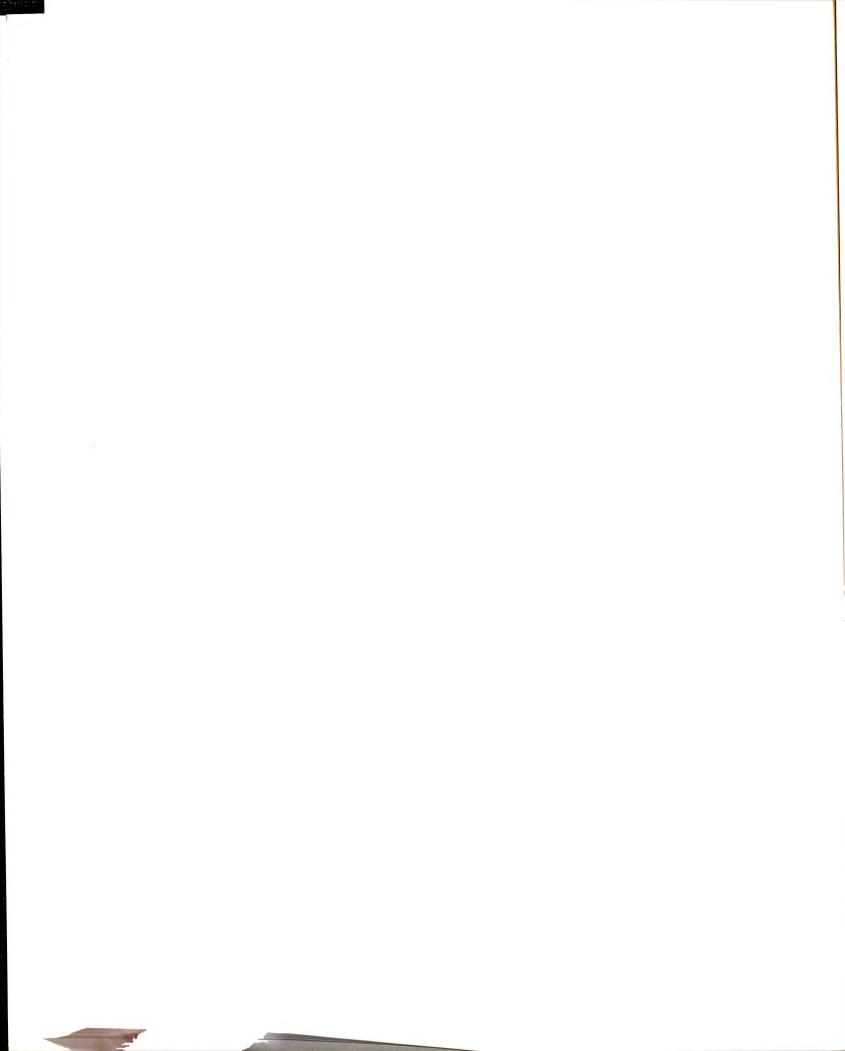
Comparison of Consistency between Intuitive and Probabilistic Judgment

In this section, we ask the question: Which mode of judgment is more susceptible to change under different problem contexts? This question will be answered globally



across all the problems and specifically for each theoretical meaningful pair of problems.

To answer the question globally, the subjects were divided into the subgroups of intuitive vs. probabilistic types according to the mode of judgment the subjects used in problem A. Then for each subgroup, the Cochran's test for related observations (Conover, 1980, p.199) was applied to test for the omnibus treatment effect of problem contexts on judgment for the remaining five problems. If the treatment effect is found to be significant, it means that the particular subgroup of subjects have significantly changed their judgment among the other five problems. This indicates that these subjects would significantly change their strategy of judgment under the influence of some problem contexts. For the probabilistic subjects, operationally defined, the treatment effect was just marginally significant with a Chi-square of 9.49, $df=4$, $p=0.050$. However, the corresponding Chi-square for the intuitive subgroup was 18.7, $df=4$, $p=0.0009$. The p values also act as a measure of the strength of the treatment effect here. The results indicated that the treatment effect of the problem contexts for the intuitive subgroup seemed to be stronger than the probabilistic subgroup, and therefore the intuitive subgroup changed more. That is to say,



probabilistic subjects tended to apply the same strategy across all problem contexts while the intuitive subjects varied their strategies when facing the different problem contexts.

The above global difference between the strategies of the intuitive and probabilistic people can also be examined specifically for those theoretically meaningful pairwise comparisons which had a significant treatment effect. The distribution of subjects for those pairwise comparisons with a significant effect on mode shifts were organized in the following table. The usual Chi-square test for no association was applied. A significant Chi-square would mean that the proportion of subjects who changed from one problem to another was dependent on the subjects' initial mode of judgment, i.e. intuitive or probabilistic strategies. In other words, the proportion of changers were different for the two mode of judgment (Bishop, Fienberg and Holland, 1988).

Three out of the four comparisons indicated a significant chi-square value which rejected equal consistency pattern for the two judgmental modes. This seems to reveal that when there was a change of judgment between two problems, the proportion of intuitive people who changed

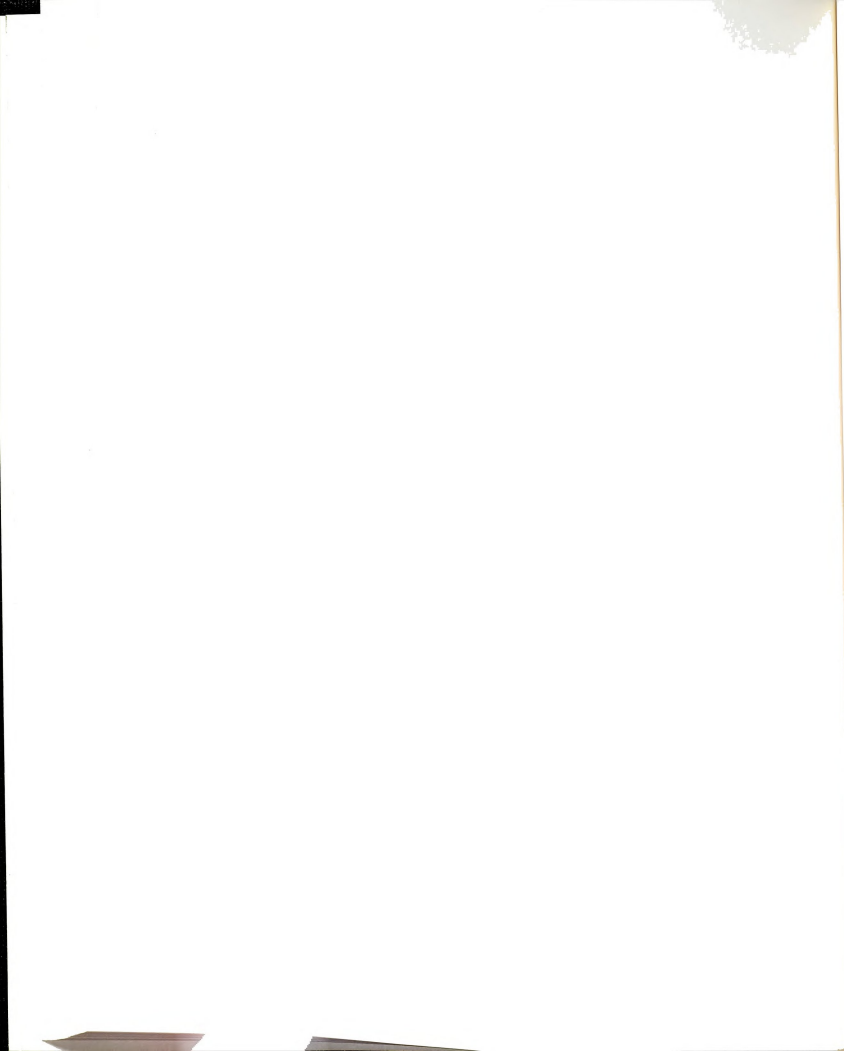


Table 10

Relative Consistency of Judgmental Mode under Various Comparisons

		(D)		Chi Square
		Same	Different	
	Intuit.	17	11	6.02
(C)	Prob.	11	0	($p < .02$)*

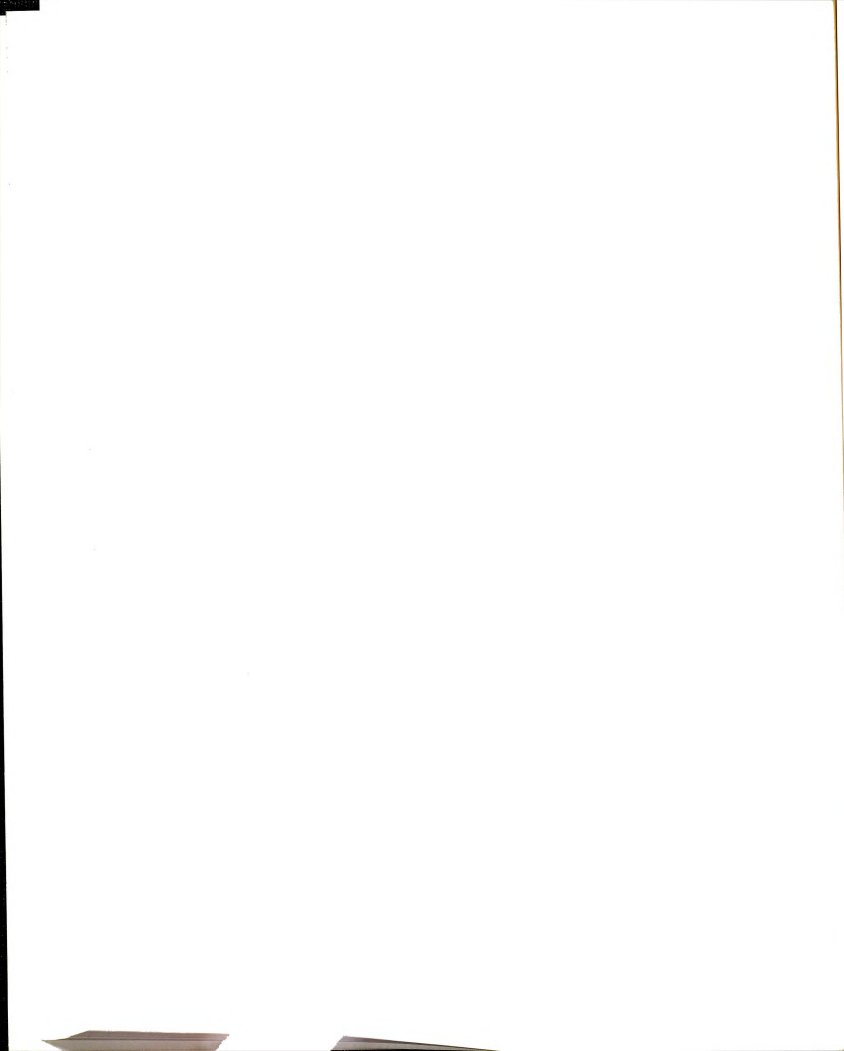
		(E)		Chi Square
		Same	Different	
	Intuit.	22	6	3.03
(C)	Prob.	12	0	($p > .05$)

		(F)		Chi Square
		Same	Different	
	Intuit.	12	13	4.8
(A)	Prob.	11	2	($p < .05$)

		(F)		Chi Square
		Same	Different	
	Intuit.	11	13	3.89
(B)	Prob.	11	3	($p < .05$)

* individual p values might change slightly due to the total number of comparisons made

judgment was higher than the that of the probabilistic people. The specific results were consistent with the global results.



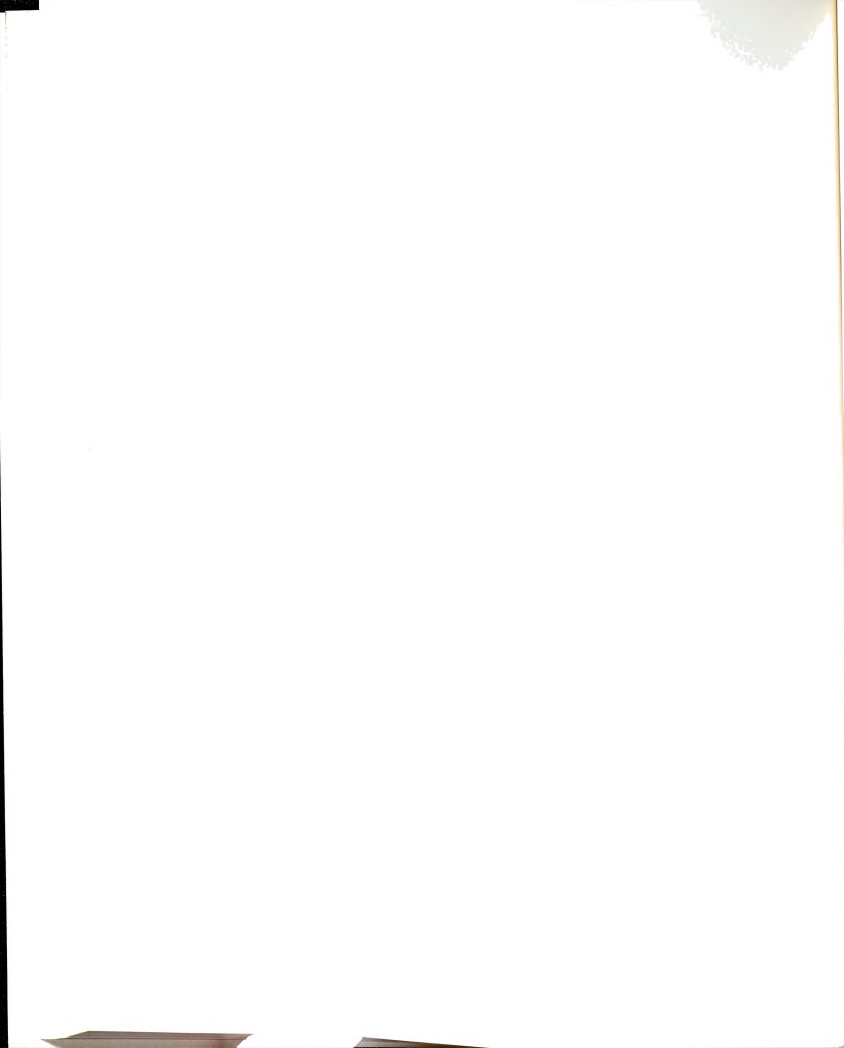
Effects of Culture, Sex, Age, Major and Bayes' Knowledge on Judgment

From our results in the chapter on the traditional analysis of subjects' numerical response, we found that the results repeated to a high degree of the research formerly done on the American and Israelian subjects. There is no reason to suspect that the Chinese subjects' responses are highly different from the pattern of responses in the West.

Our subjects were coded to belong to either arts and social science or natural science and engineering. They were also asked whether they learned Bayes' rule before.

Crosstabulations between major area and mode of judgment obtained no significant chi-square for test of independence for all problems used in this study. Thus major area of study does not seem to affect whether people use intuitive or probabilistic mode of judgment. Chi-square tests for the effect of knowledge about Bayes' rule also obtained no significant results. In our sample of subjects all of whom knew at least some simple probability theories, knowing Bayes' rule is probably an indicator of better statistical knowledge. Our results showed that judgmental mode was not affected by better statistical knowledge at all.

Crosstabulation between sex and mode of judgment obtained significant chi-square for problem F only, chi-



square was 4.51, $p=0.0337$. Our female subjects seemed to view the problem about the machine vision with a more intuitive perspective than the males. However, in general the sex effect is not dominant.

Table 11

Effects of Sex, Major and Knowledge about Bayes' rule on Judgment

Chi-square statistics with corresponding probability value shown in parentheses.

Problem		Sex		Major		Knowledge of
N						Bayes' Rule
A	39	0.300	(.584)	0.014	(.905)	0.551 (.458)
B	39	0.0410	(.839)	0.742	(.389)	2.839 (.092)
C	40	1.91	(.168)	0.476	(.490)	0.0770 (.781)
D	39	0.0332	(.855)	2.17	(.140)	0.300 (.584)
E	40	1.62	(.204)	0.404	(.525)	0.331 (.565)
F	39	4.51*	(.0337)	0.0144	(.905)	0.365 (.546)

Note * : Significant at 0.05 level

Correlations between the subject's age and his mode of judgment across all problems were also computed. None of them were significant. The correlation coefficients are tabulated below:

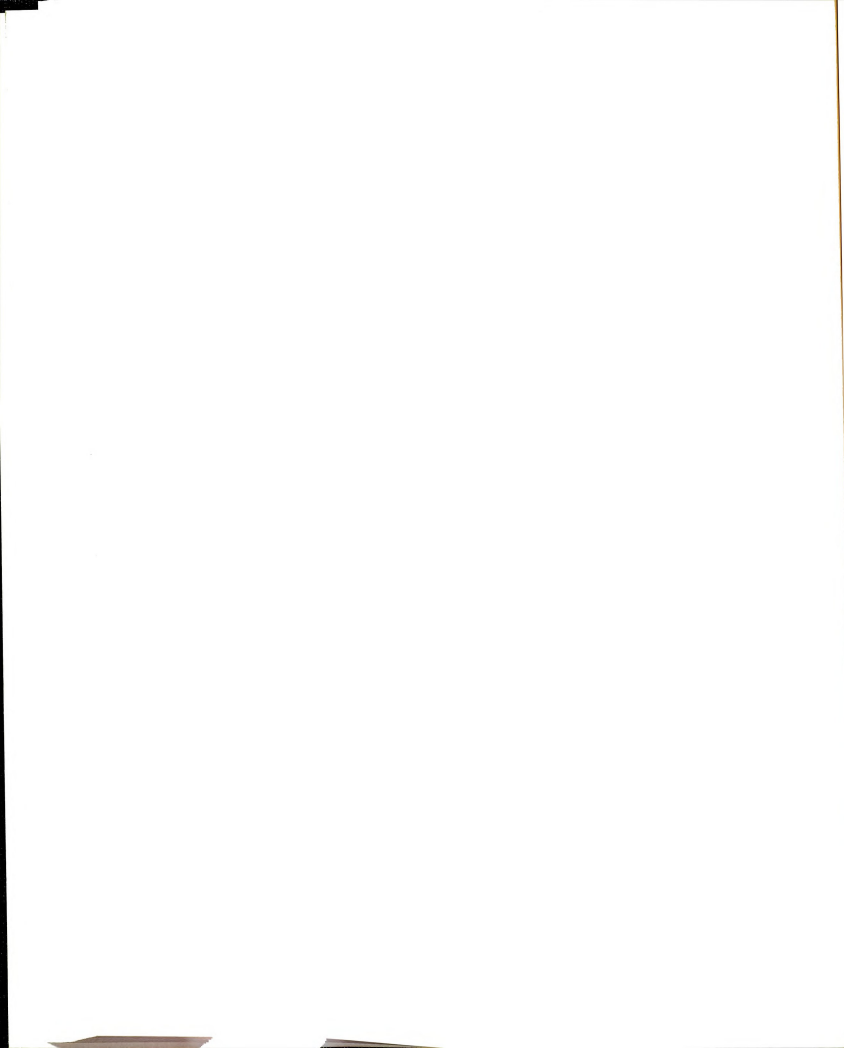
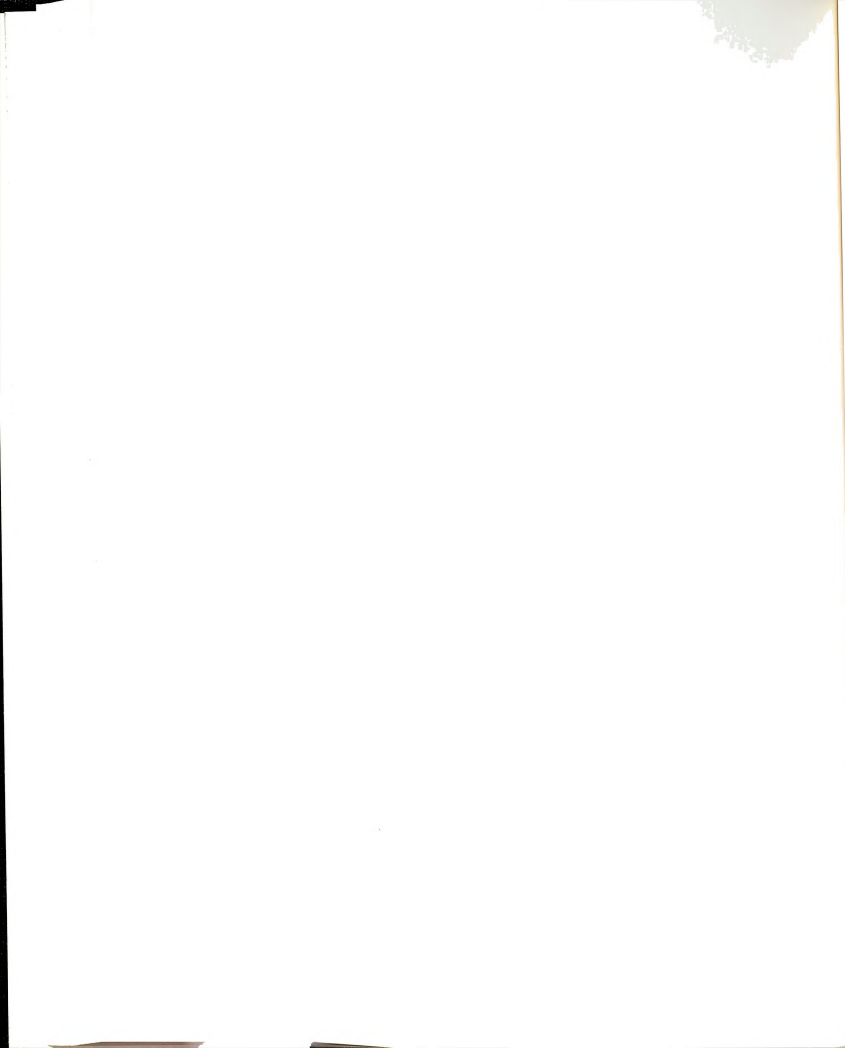


Table 12

Correlation between age and mode of judgment

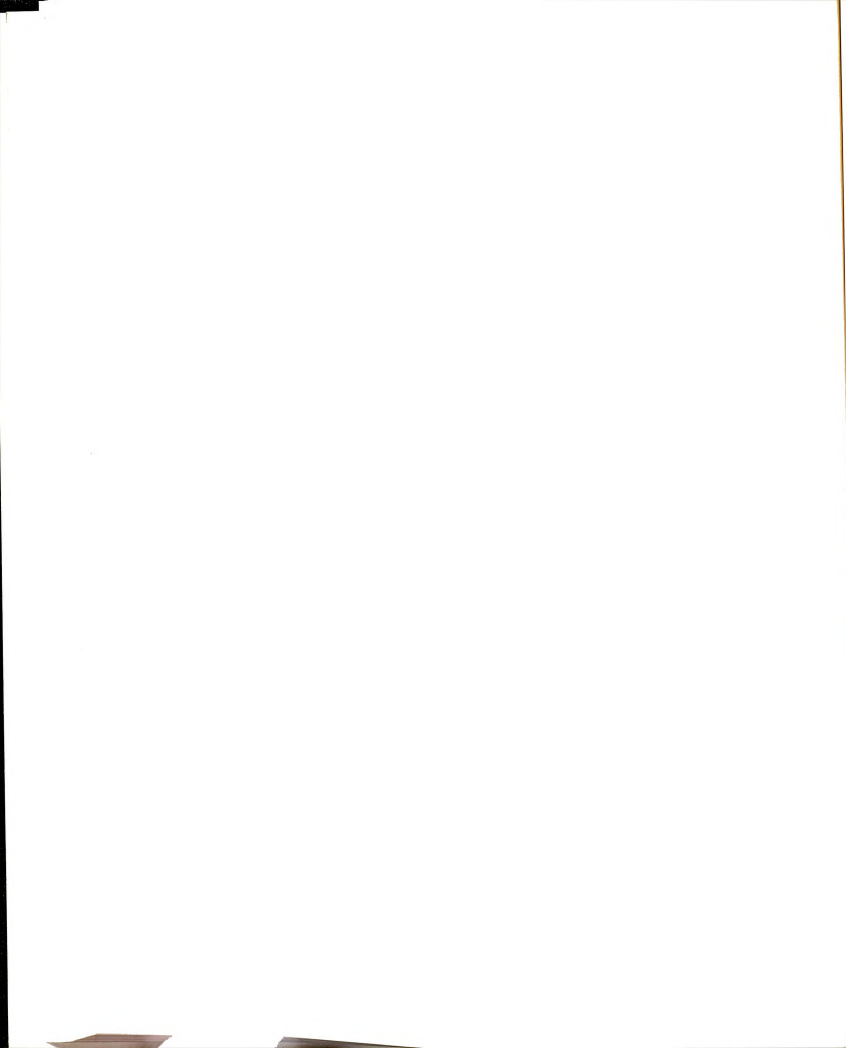
Problem	Correlation	N	p
A	-0.0446	37	0.794
B	-0.1949	37	0.248
C	0.0219	38	0.896
D	0.2058	37	0.222
E	-0.0520	38	0.756
F	-0.0466	37	0.784



DISCUSSIONS

Histograms of our subjects' responses confirmed the bi-modal or multi-modal nature of the response from the base rate problems. The usual analysis based on the mean or median is called into question. Our study demonstrated that intuitive and probabilistic mode of judgment can be successfully delineated in the base rate problem. Probabilistic mode of judgment conforms the calculus of chance, or the "urn" model and involves explicit application of the axiomatic probability theory like the multiplicative rule or the conditional probability rule. Intuitive mode of judgment do not rely on abstract axiomatic theory. Relevance, importance and weights of the evidence are also considered. Logical deduction and narratives are also used to fill in the gaps of the given information.

There does not seem to be any difference in numerical responses between our Chinese subjects and the American or Israelian subjects of some earlier research. Whether someone is in arts or science does not seem to affect what mode of judgement is used. Sex has a small effect. In 2 out of 6 problems, more female subjects appear to be judging intuitively than males.

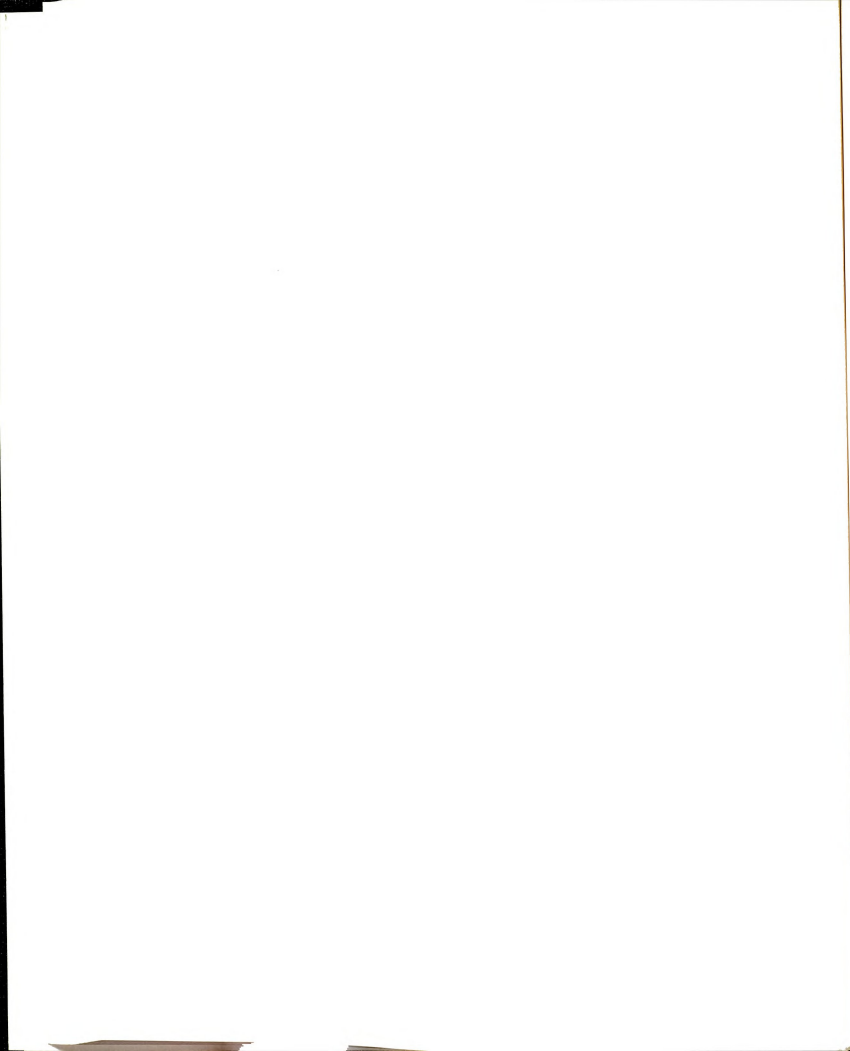


The context of base rate problem does seem to affect the distribution of the mode of judgment in our sample of subjects. The distribution of subjects turn towards being more probabilistic under the following problem contexts:

1. Diagnostic information with low diagnosticity for judgment.
2. Extreme low or high base rate.
3. A pure physical or mechanical context.

While the third result is a new discovery, results 2 and 3 are parallel to former research about factors affecting the use of base rate in social judgment.

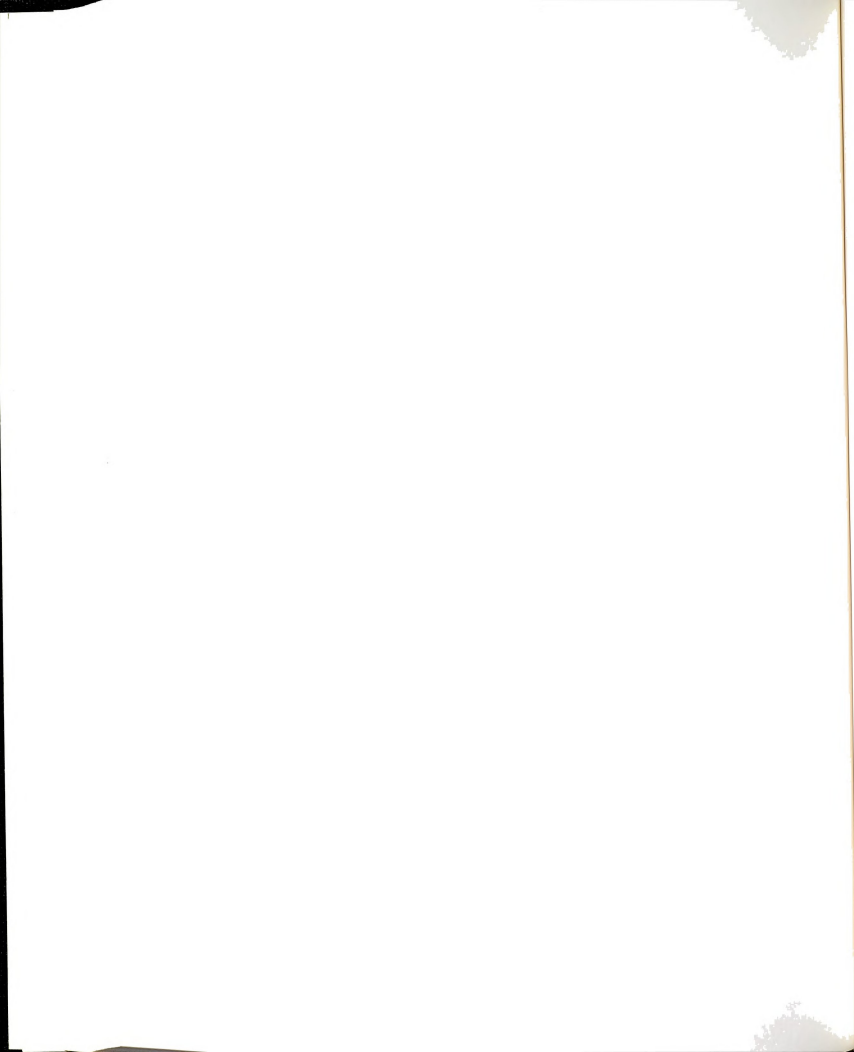
There is no significant difference between the distributions of judgment of the cab problem with a causal base rate and that with a non-causal base rate. Former research (Tversky and Kahneman, 1980) used to claim that more subjects used the base rate under the causal condition because the median response under the causal condition (60%) was closer to the base rate (15%) than that under the non-causal condition (80%). From our data, it appears that although 5 out of 25 subjects did lower their response under a causal condition which decrease the median response to 70%, they were still using intuitive judgment. Tested from the standpoint of judgmental mode, only 3 subjects changed mode and the result was not statistically significant.



Chi-square tests of independence and the Cochran's test for related observation revealed that probabilistic judges are less susceptible to judgmental mode shift than the intuitive subjects. It appeared that probabilistic judges just plug in the numbers by some probability rule, although incorrect, and remain relatively unchanged by the experimental manipulations. Intuitive subjects would consider the experimental information and shift to the probabilistic mode as regarded necessary.

Qualitative Aspects of Intuitive Judgment

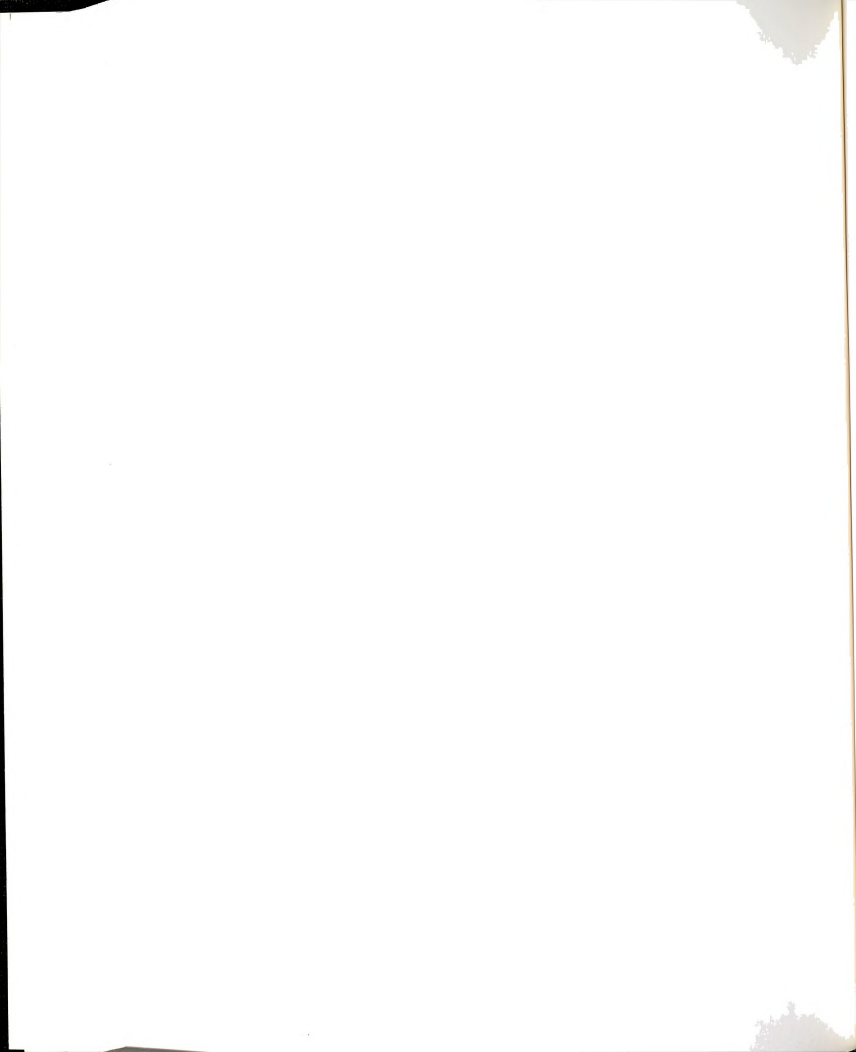
The probabilistic mode of judgment which complies with the calculus of probability is not very interesting per se, at least not as interesting as the functioning mechanism of the intuitive mode of judgment. The reason is a high proportion of subjects which ranged from 35% to 70%, exhibited intuitive mode of judgment in the sample of problems we tested. Many interesting questions could arise, such as whether people who make intuitive judgment are irrational, or is it just a matter of education? How do politicians, physicians and bankers conduct their business, presumably through intuitive judgment? Can we trust our judge and juries if they are intuitive thinkers? Alternatively, besides knowing its "errors", can we learn



anything worthwhile from intuitive judgment? Could artificial intelligence learn something from intuitive judgment? These list can go on, indefinitely. Having gone through the experience of interviewing 40 talented and educated people about how they make judgments on those base rate problems, I try to present my opinions, if only partial answers, to the above important questions.

A theoretical world is like a base rate problem. The problem consists of two and only two piece of important information: the base rate and the diagnostic information. There is a theoretical optimum, obtained by applying Bayes' formula to these two information. Numerical answer is exact or up to the number of decimal points we desire. The theoretical world abstracts the real world and is therefore not the real world. Depending on the quality of such abstraction, the theoretical world represents the real world variably.

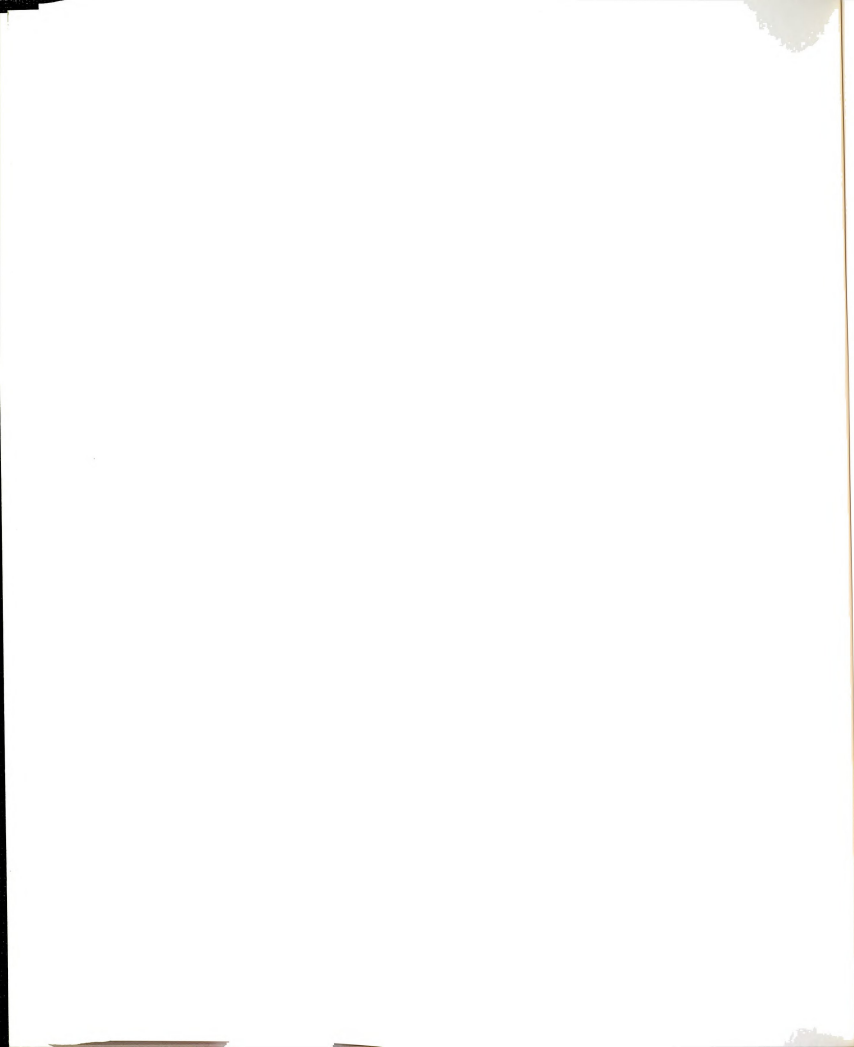
A real world is like the world we live in. We can doubt. We can ask question and are given some answers. We can challenge the authority. We can find out more about something if we are not sure. There are established rules to do certain things. We perceive quality; something is well done and some are not. We know that people tell lies; we do not believe in everything.



Both intuitive and probabilistic mode of judgment belong to the real world. The intuitive mode is with us all the time and the probabilistic mode exists only when we are making abstractions or theorizing.

The base rate problem with its Bayes' solution is like a faultless world. There are absolutely no doubt or problem about anything, except that you are supposed to make a probabilistic estimate of the diagnostic information (e.g. given the described characteristics of the person, how much is the probability that this person is an engineer?). Then this probability estimate together with the base rate are supposed to be entered into the Bayes' formula to obtain the optimal solution to the problem. The intuitive subjects do not work like this. They do not take the information for granted. They appear to function in a complex realm, as if in the real world. The following discussion, supported by the subjects' , are used as examples to highlight the special cognitive and meta-cognitive aspects of the intuitive mode of judgment. The conditions under which intuitive mode of judgment would err are also discussed.

Intuitive judges do not take the information for granted. People in an intuitive mode of judgment do not take things for granted. You cannot easily get them to obey by saying: "Forget about everything else, just give me an

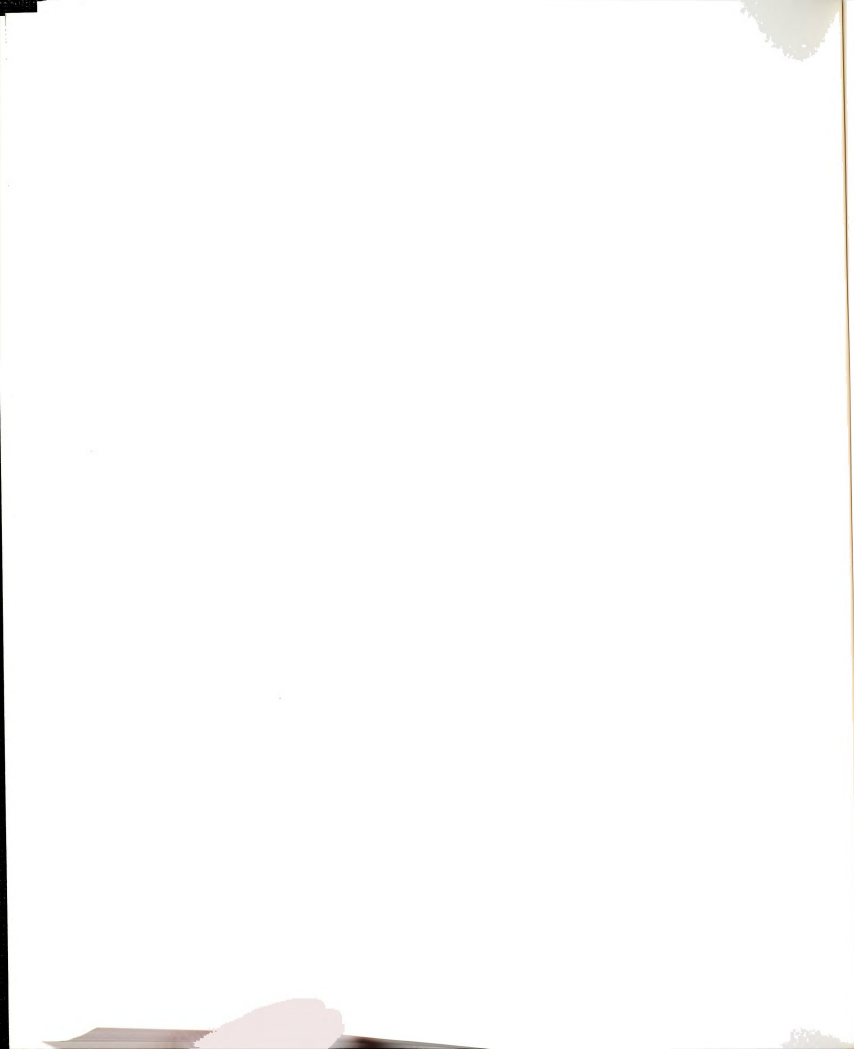


answer by looking at the two given key information." Because that is not the usual way judgment holds in the real world. These people challenge the logic implied by the question. For example by saying: "More cars don't mean they must crash more." We all know the simple fact that a student who studies longer hours might not do better than the brighter student who studies less. A person is generally considered of less intelligent if he or she can only follow what he or she is told and question nothing. Relevant excerpts from subjects' reasonings are listed below. (S and A stand for subject and problem number respectively; the numeral in between is the subject number)

1. Many (more) cars don't mean they must crash (more).
(S31A)
2. A car was involved ... Two cab companies ... It doesn't mean that the car involved in the accident is a cab! (S8A)
3. When the accident happens, maybe green cars are not around the scene. (S23A)
4. Maybe blue cab's business is better. (S29A)

Intuitive judges look for more relevant information.

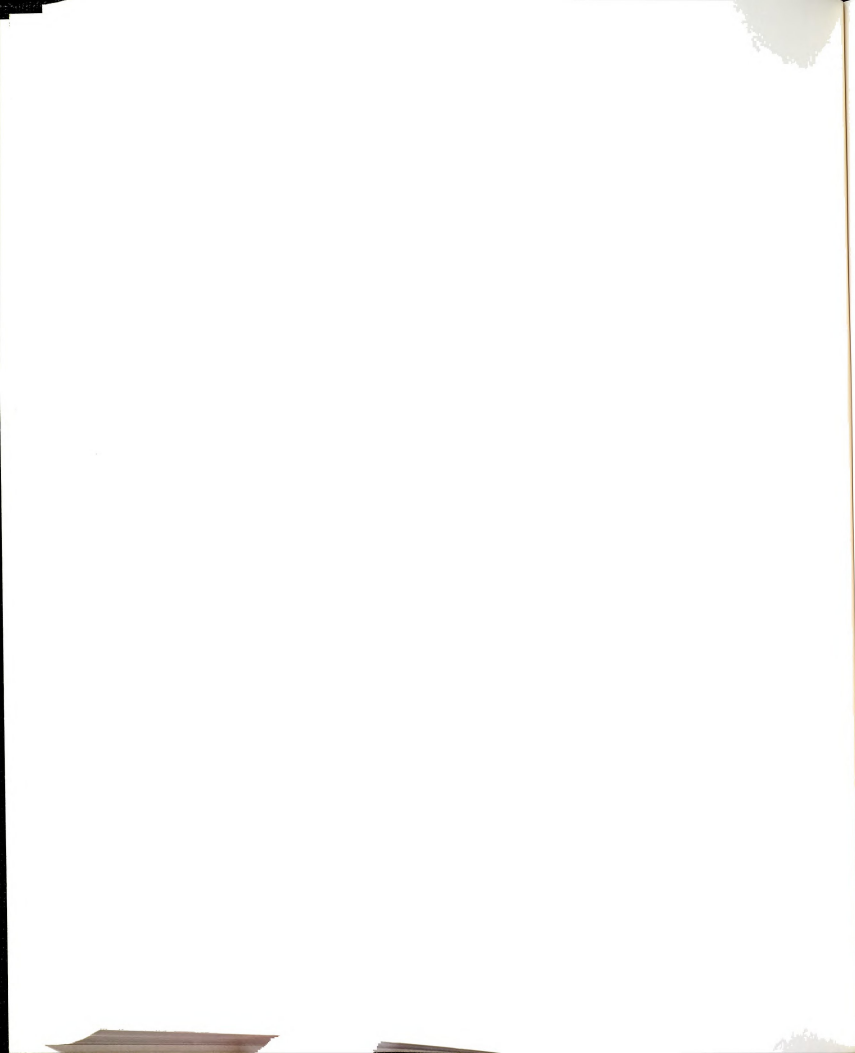
There is evidence that the intuitive subject tries to obtain all the important information relevant to the problem he wants to judge just as what he will do to a similar



situation in the real world. Of course, he cannot do so in an experiment. He must supply his own assumptions, drawn from his experience. There is a dialogue between a person and his world in the real world. He can always ask for more information or search it by his own. No such thing exist in an experiment. Experimental results might possibly differ in an imagined experiment which would supply any additional information the subject wants. Without the information considered crucial in judging a given case, the subject can only make his own assumptions or quit if he is permitted to do so. (Note: Some subjects questioned me for additional information which I did not have and certain subjects expressed to me that it was difficult for them to judge without knowing more about the case) Evans (1989) missed this point by considering the additional information that subjects made up led them towards wrong judgment (against the norm in the theoretical world). Some examples are as follows:

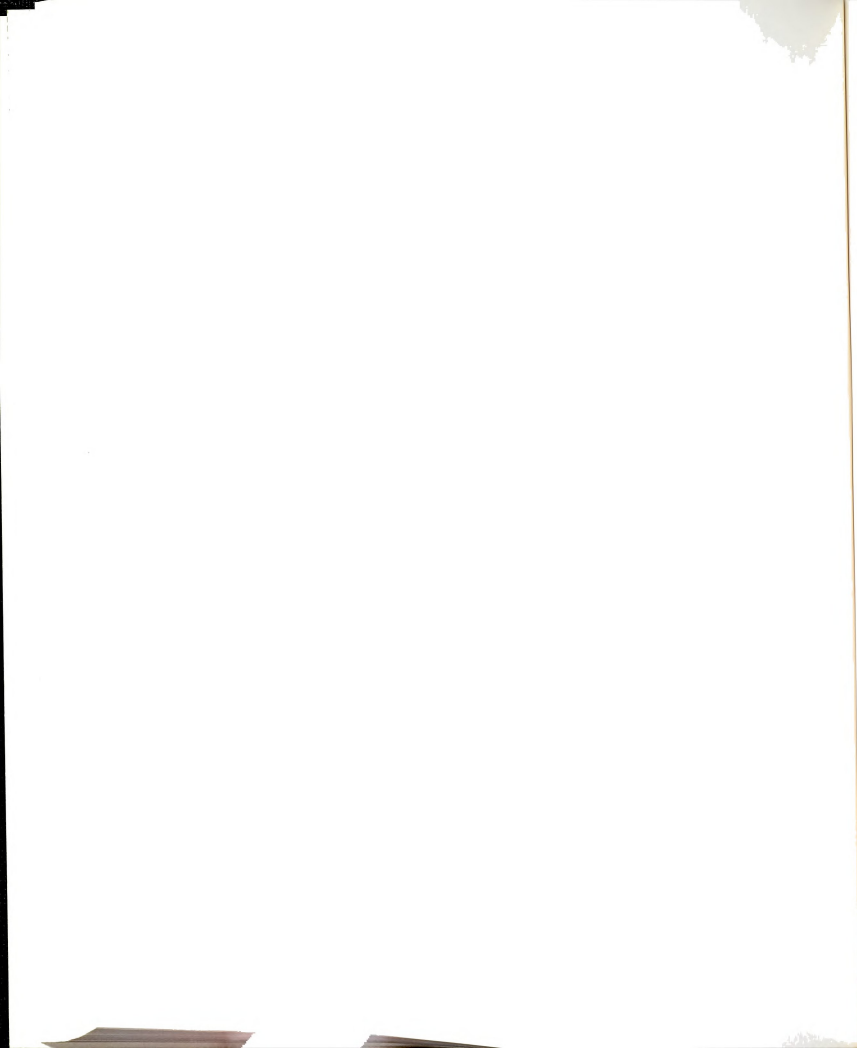
1. Other relevant factors: driver, car's machinery. May be blue higher (better) than green. (S22A)
2. Other similar data (about the witness) might be proposed (needed). (S21A)

Intuitive subjects supply their own knowledge and assumptions. When the subject has to supply his own



assumptions to fill in gaps of the problem to be able to judge, he draws his assumptions from his repertoire of knowledge and belief useful to the given situation. For example, our subjects think what is like for a car accident, for a court investigation or for the personality of some engineers they know or have heard of. This information is life-like and thus usually comprises many factors and dimensions. As a result, these assumptions should in general exceed or contradict the assumptions that the problem intends. Nevertheless, these assumptions of knowledge and belief might be wrong and inaccurate by themselves, or they might be wrongly applied to the given situation. This is an example of judgmental error in the real world. Some examples from our subjects' reasonings are as follows:

1. Lawyer is not liked by the colleagues. (S7D)
2. Engineer works alone. (S40D)
3. Engineer work independently. (S34D)
4. Engineer work in team. (S14D)
5. Lawyer is very competitive, unlikely be liked by colleagues. (S20D)
6. Lawyer like this cannot be successful (S9C)
7. An engineer is freer to discover and is allowed to have mistakes. (S23D)

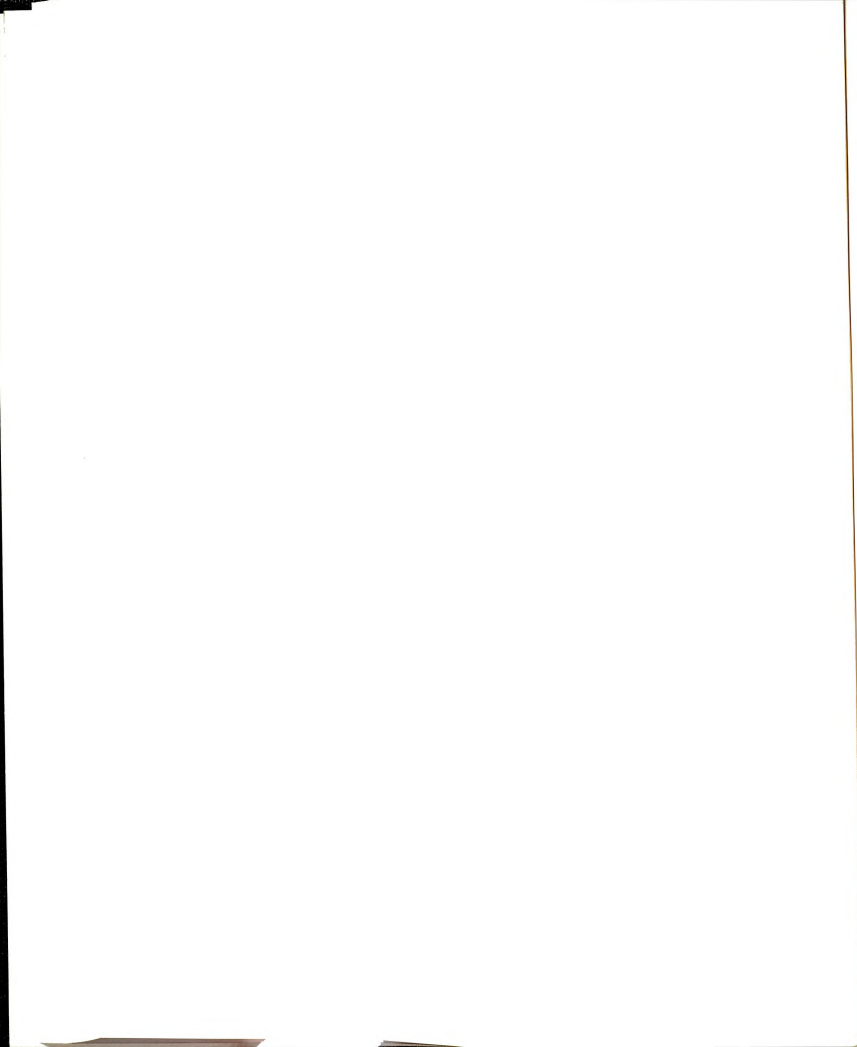


8. Aged 45, a lawyer and have no interest in social and political affairs is unlikely. (S29E)
9. Since there are fewer blue cars (in the city), witness will pay more attention if it (the involved car) is blue. Percentage (the witness's ability in recognizing colors) should be higher than 80% (the given). (S33A)

In real world resource of time and material are limited for any given person for a certain purpose. One cannot obtain any information one wants. One must plan to obtain the most important information relevant to the situation in the most economical way. Of course, one cannot always do this optimally. One is then stuck with information with second class value or miss the chance of getting the information at all. This may be why people err in the real world.

Intuitive judges balance information by its relevance.

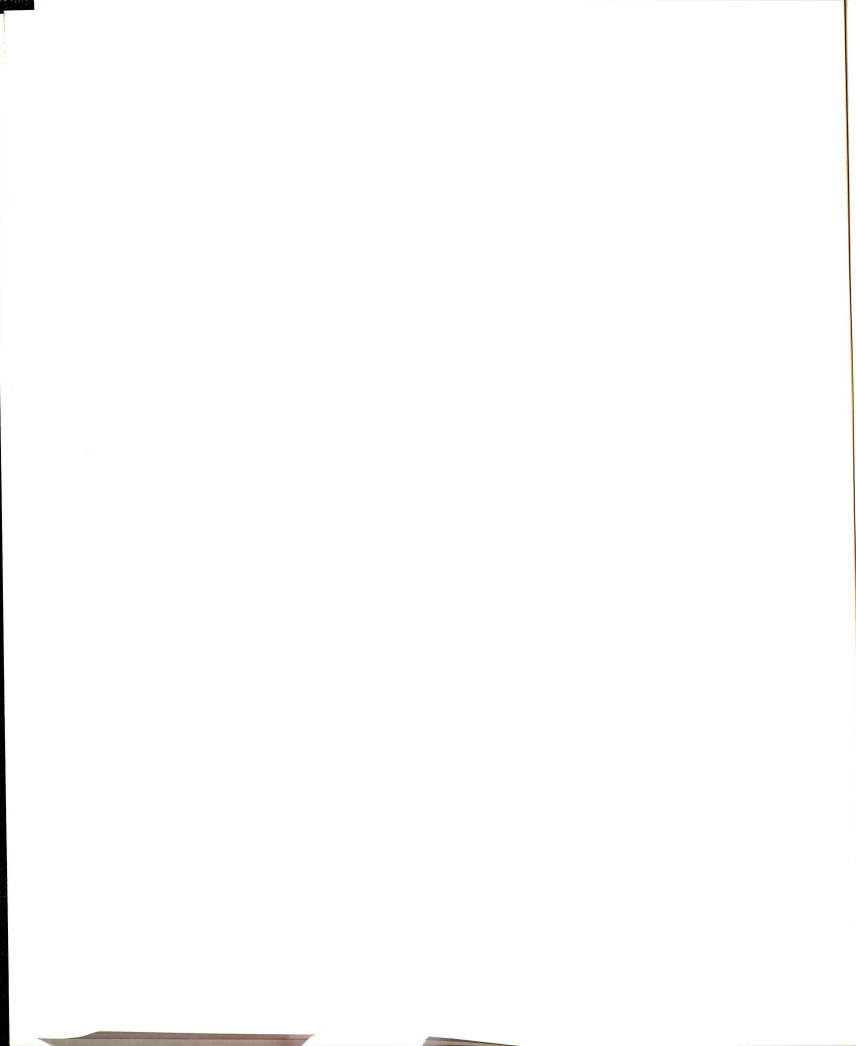
People in real world consider the importance, value and relevance of things. Our subjects look at information as if weights are attached to them. They decide subjectively what is relevant, what is important. They can compare the importance of any one piece of information in the given case relevant to finding the solution, not just two piece of information. the experimenter intends.



A person functioning in the real world suffers from limited cognitive ability. Our brain cannot recall or compute like a computer. Our mind works mostly in discrete levels; seldom on a continuum. For example, our study shows that intuitive people can become probabilistic when the base rate is made to be very extreme, like from 30% to 1%. The Bayes' rule can give an answer no matter how slight a change is the base rate. But a person can only respond when the change is subjectively detectable and is being felt as of a significant magnitude. Moreover, when the situation is complex and of numerous dimensions, a person might not be able to summarize all the information to a level that he can manage cognitively. He is then bound to make error in judgment.

Excerpts which demonstrates the importance of relevance are:

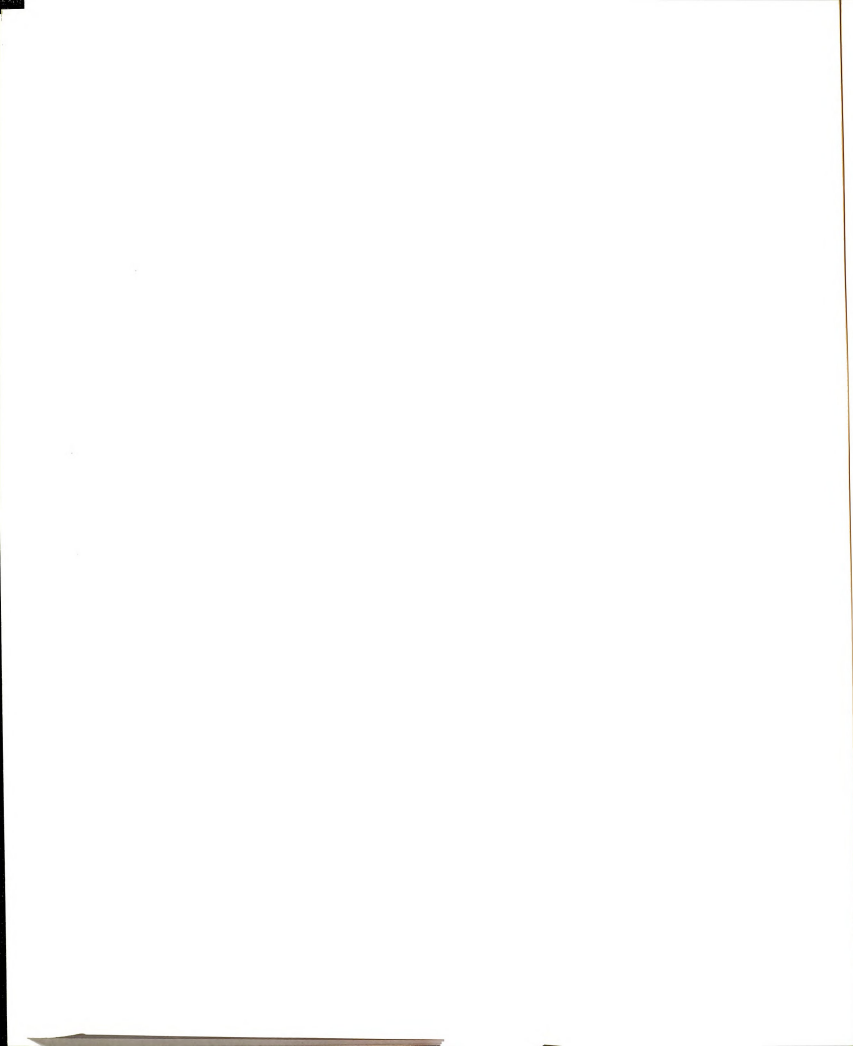
1. I focus on the character. (S7E)
2. Just like the taxi problem, ratio does not have much meaning. (S40E)
3. It has relation with the character, not number. (S22E)
4. I based my judgment on the description, statistical data has no relation. (S6E, S31E)
5. The experimental evidence is primary, the frequency data is only secondary. (S7A,B)
6. It's a single event, witness is more important. (S24A)



7. I trust repeated experiment. (S39A)
8. I trust witness, the rate is irrelevant. (N=11,A)
9. The past and the present are independent. (N=17,B)
10. I judge according to the error percentage, ratio has no influence. (N=14,F)
11. The description is vague, it can be either an engineer or a lawyer. (S6,19,22,21,28,D)

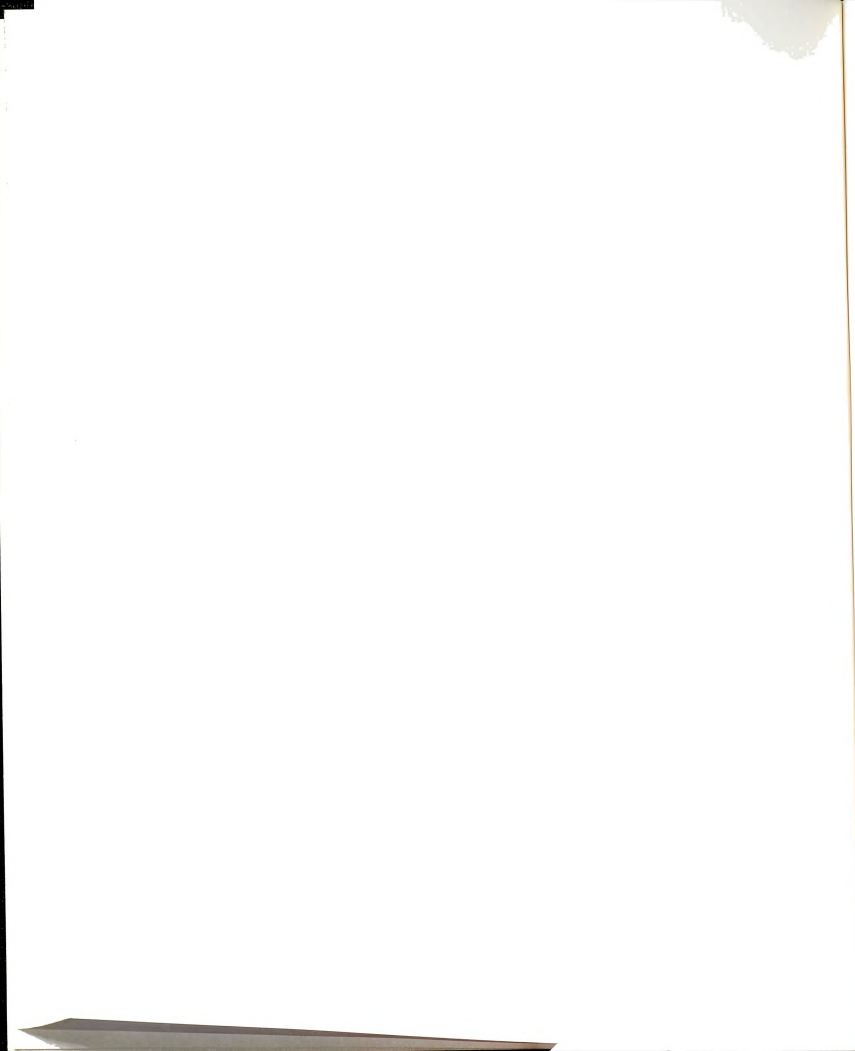
Excerpts which illustrate the intuitive subject's use of balancing information are:

1. Lower witness's rate, since the two rates are very different now. (S27A,S30A)
2. It has a higher accident rate, we should lower the witness's reliability. (S33,26,1,28,30,B)
3. The majority is circle. Lower the error percentage. (S27F)
4. The extreme base rate has effect, but it is like an engineer. It should be higher than 1 %. (S19,21,34,14,27E)
5. It is an engineer. Since there is only one engineer, we should lower the probability. (S24,39,30,1,28,33,20E)
6. We are not given any (useful) information, we have to depend on the earlier data. (S9D)



Implications for the Debate about Human Rationality

Research in human judgment of the last decade had been a debating ground for the forum about human rationality (e.g. Cohen 1981). The proposal that humans are not as rational as they might seem came from experimenters who recently discovered that people fail to meet the norms in many deductive and inferential tasks (e.g. Evans, 1989; Kahneman & Tversky 1982; Nisbett & Ross, 1980). The defense for human rationality were usually of a theoretical nature. For example, Cohen (1981) presented the views that some subjects might be functioning in some other equally valid concepts of probability, one of which belongs to the Pascalian probability elaborated in Cohen (1977). But the question is if the other type of probability is equally valid, why don't subjects reach answers close to the mathematical or statistical norm? White (1984) suggested that practical judgment is concrete, everyday and unstructured and the objective of judgement is not to produce an outcome that is right in a normative sense but an outcome that satisfies the practical concerns. Still this do not explain very well why people fail in experiments about judgment and why mathematical and statistical norms can be applicable. There are data which support the experimenters while the theorists usually present no empirical support. Because of this reason, the debate was not well settled. I



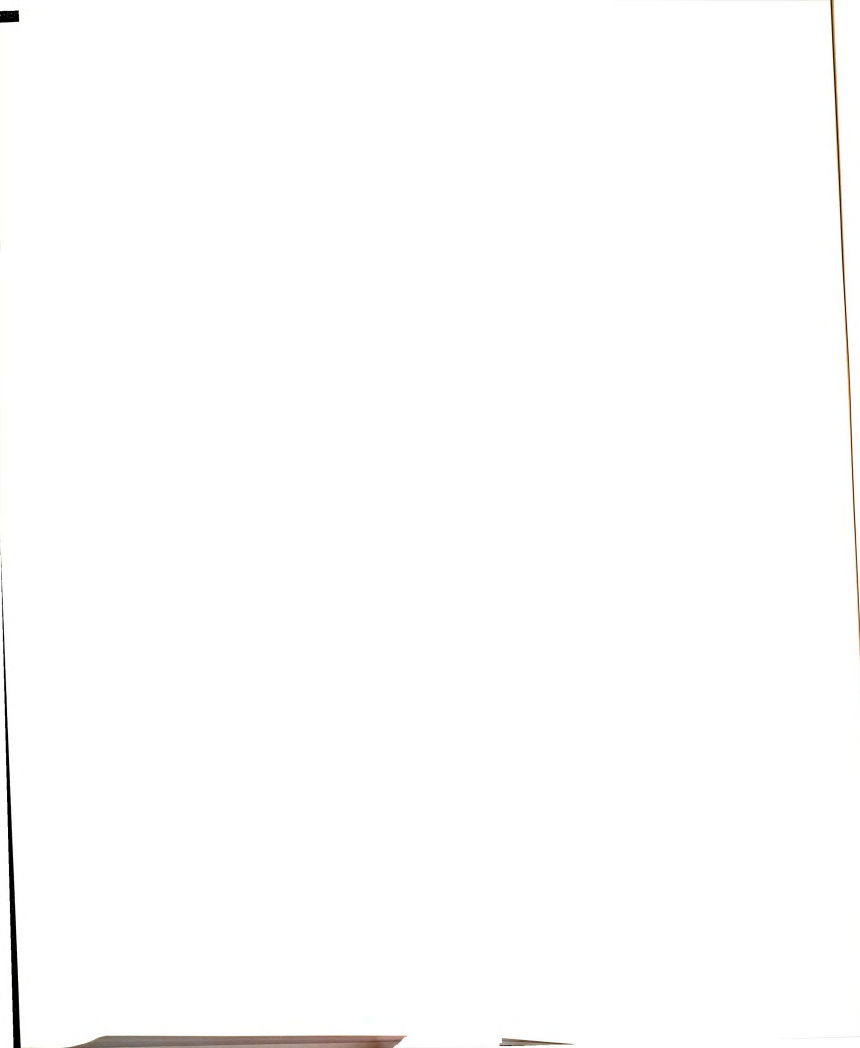
would present a different view of the debate, based on the qualitative and quantitative study that I conducted.

The Majority of people cannot reach the exact answer to a mathematical or statistical problem. That is for sure. It take years of education to learn the higher level mathematics. In my experiment on 40 Chinese students, mostly graduate student; only 22 (55%) claimed that they have learned about Bayes' theorem before. However, none of them remembered the formula off hand. Only three subjects in my sample used conditional probability to reach the same answer required by Bayes' rule in five instances. Well if most people cannot reach the mathematical norm, doesn't that mean they are not rational enough, to function in the modern world full of uncertainty?

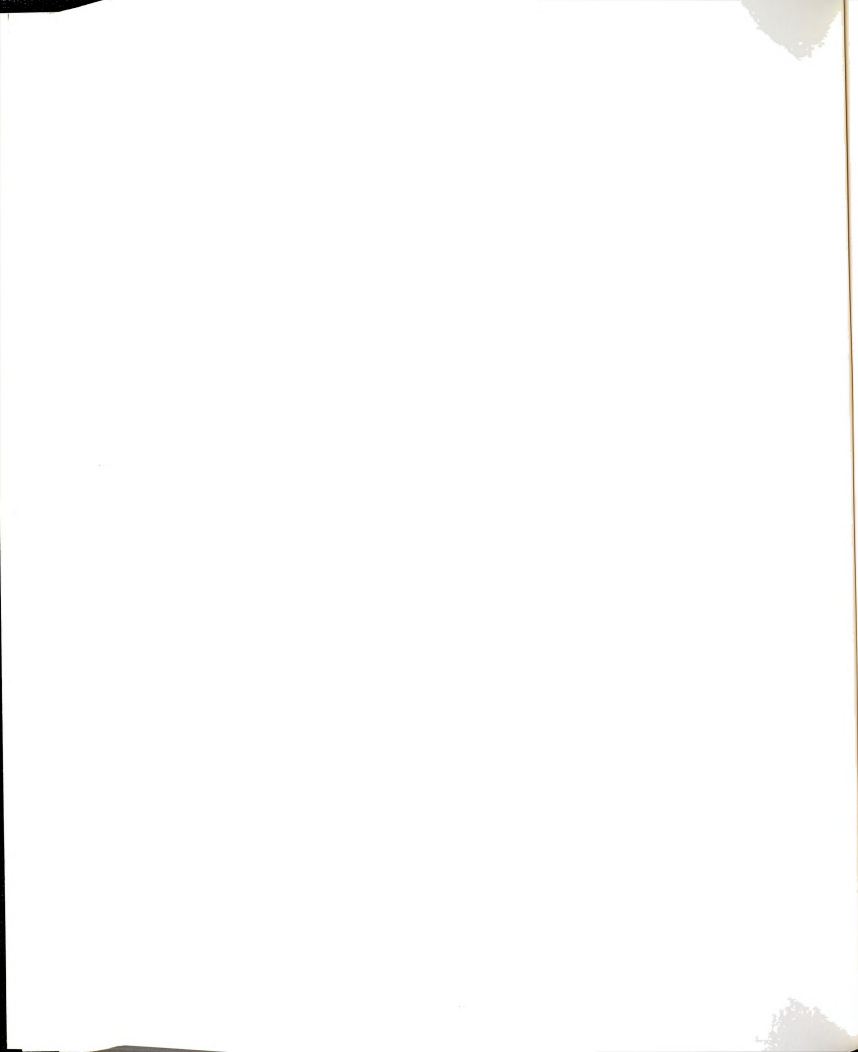
With the insight from my experiment, I would answer that if people always automatically and mechanically apply Bayes' theorem in their judgment, that will show only that these people lack the normal cognitive and meta-cognitive intelligence for judgment in the real world. As I have discussed before, the real world phenomena as presented to us is often incomplete, untrue or lack of relevant information unless we search for it. Problems in real world are always multi-dimensional. Information has levels of quality. The real world is not the faultless world as in the Bayes' problem that researchers intend to test their

subjects. Our study documented that subjects do not take "evidence" for granted. They question authoritative judgment by asking questions, ask for more information or search for additional information himself if considered necessary and possible. The subjects have past knowledge and beliefs that they would bring the situation to make the best judgment. They weigh the evidence by considering its relevance, credibility and importance. Although in a limited cognitive capacity, they summarize information and balance the weights of the information. They also uses logic to examine the quality, weights and credibility of the evidence.

These characteristics of intuitive judgment, empirically supported by our data, are quality that is unmatched by a simple-minded and mechanical application of a mathematical or statistical formula. The reason is again that real world problems are generally complex, incomplete and non-routine. For example, our subjects had reasons to question whether besides the quantities of cabs, other things were equal. They questioned whether the business, machinery, driver training between the two cab companies were the same. These are important factors which could affect the decision making. Without these information, the intuitive subjects refused to accept cab sizes as the appropriate signal for the likelihoods of making accidents.



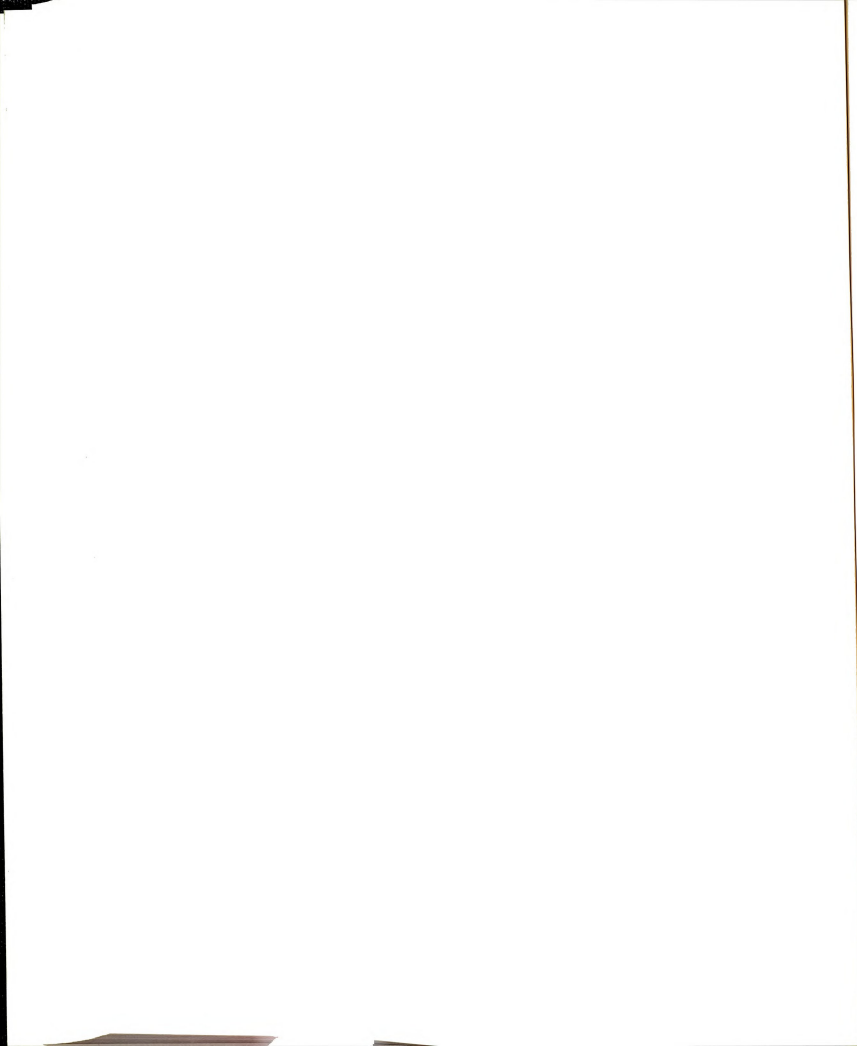
A mechanical application of a mathematical rule would not take care of other possible important but unconsidered factors. Without the crucial information, a sensible person would either search for additional information or supply his own assumptions and beliefs. However, simple application of some mathematical norm could just work on the given limited data which might be irrelevant or wrong. An intuitive subject can decide whether a piece of information is relevant or not, and he or she can weigh the evidence by its relevance. Obviously, plugging in some pre-determined formula simple-mindedly would not consider the weights of information in a sensible way. The reason that subjects fail to meet the mathematical norm in experiments is because these subjects are employing intuitive judgment, as if what they would do in the real world. They employed cognitive and meta-cognitive skills that go beyond the information given, the boundary that the experimenters want to impose. One cannot really determine whether the subject's intuitive judgment is right or the experimenter's normative judgment is right? Because looking from the perspective of the real world, the experimenter who expects their subjects to plug in a formula without questioning the details of the given information is simple-minded. Alternatively, from the perspective of the theoretical world, the subjects are just



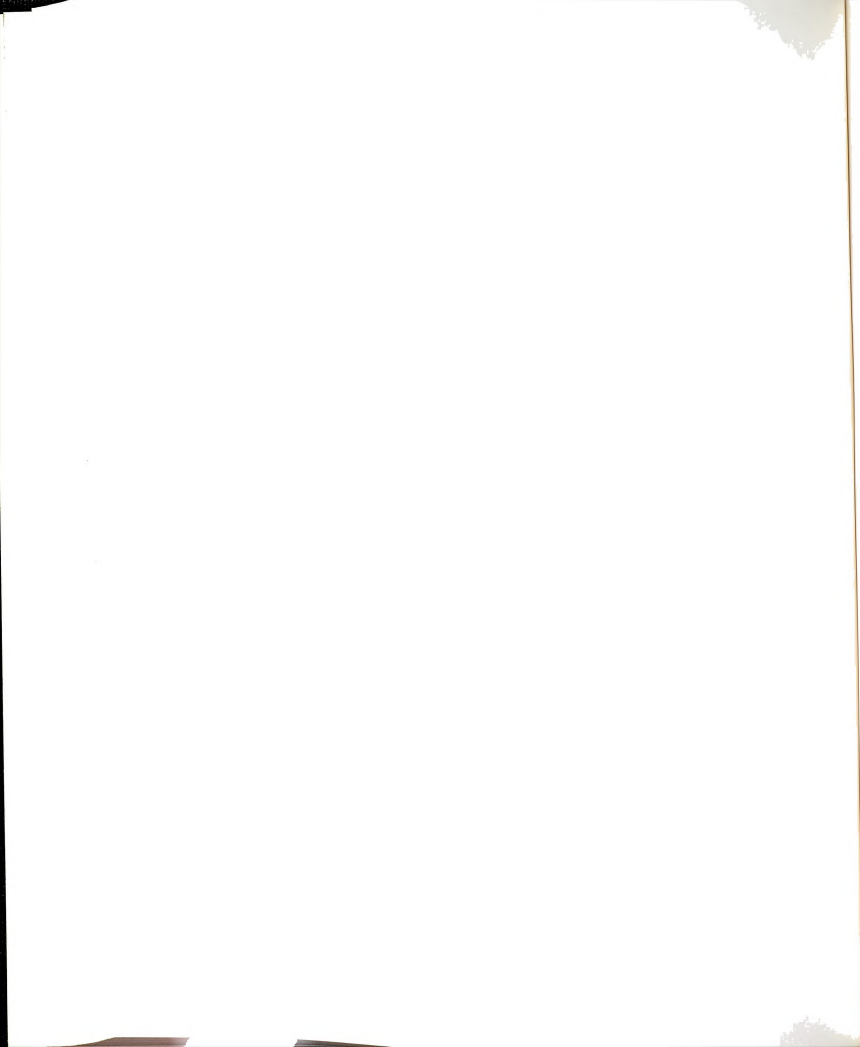
wrong, doing "unnecessary" things and bringing in "redundant" information within an experiment.

Well, does it mean that people's intuitive judgment never err? By no means. From our discussion of the last section, we suggested that people could make mistakes in most of their cognitive and meta-cognitive thinking. For example, people can be over suspicious, heading in the wrong direction for information and putting off decisions. Their beliefs and knowledge that they bring to the situation can come from a poor and distorted memory, or from a wrong interpretation. Cognitive capacity can also limit the ability to detect minute change in the environment. When the situation become more complex, people might not be able to summarize information properly and might fail to balance the weights or relevance of the information correctly.

Then what is the role of the mathematical and statistical norms for judgment in the real world? Mathematical and statistical norms can be used to enrich our knowledge and belief bases. However, these norms must not be applied mechanically to the real world without using the above-mentioned cognitive and meta-cognitive skills of intuitive judgment. Whether a mathematical or statistical formula adequately represents the real world for real world purposes must be pre-judged using intuitive judgment.



Finally, does it mean that the intuitive mind always functions better than a computer? A computer which can only do mechanical application of mathematical norms would fail to compete with the intuitive mind in many real world issues. However, the speed and storage capacity of a computer usually surpass those of a person. For designing an artificial intelligence which can make decisions in the complex real world, we must model the machine after the cognitive and meta-cognitive abilities of human's intuitive judgment as above-mentioned. Then the That is to say, the artificial intelligence should include a search for additional relevant information, going beyond the given situation and assumptions. It must be able to supply its own assumptions and knowledge when the given problem does not comply with the usual framework of analysis. It must also decide the relevance of information with respect to the problem solving. It should be able to balance a large number of information according to the relevance of each piece of information. Then the advantages of speed, storage and precision of the computer can make for the relatively limited capacity of the ordinary person. A computer can also be designed to help make judgments when the situation is becoming too complex for a person cannot summarize, weigh or balance the over-loaded information properly. Motivational problems in human judgment such as excessive emotion,



prejudice, vested interest or fatigue also give a rationale for machine decision making. However, in any case, cognitive and meta-cognitive skills of intuitive judgment has to be given a dominant position in the design of machine decision-making.

Implications for the Debate about Clinical vs. Statistical Prediction

The clinical vs. statistical prediction problem has been an important debate in psychology for the last several decades (Meehl, 1954, 1986; Holt, 1958, 1986). Because of the apparent similarities between clinical judgment and the intuitive mode of judgment, our descriptions of the characteristics of intuitive judgment can have some implications for this debate. According to the last section's discussion, clinical prediction can benefit from exercising the advantages of intuitive judgment in real life clinical judgment. In theory and possibly in practice, clinical judgment could take into account the unique characteristic pattern of information about an individual and the clinician can obtain additional relevant information if necessary. Statistical prediction is less flexible in this regard because usually the predictor variables are determined in advance and are the same for every individual.

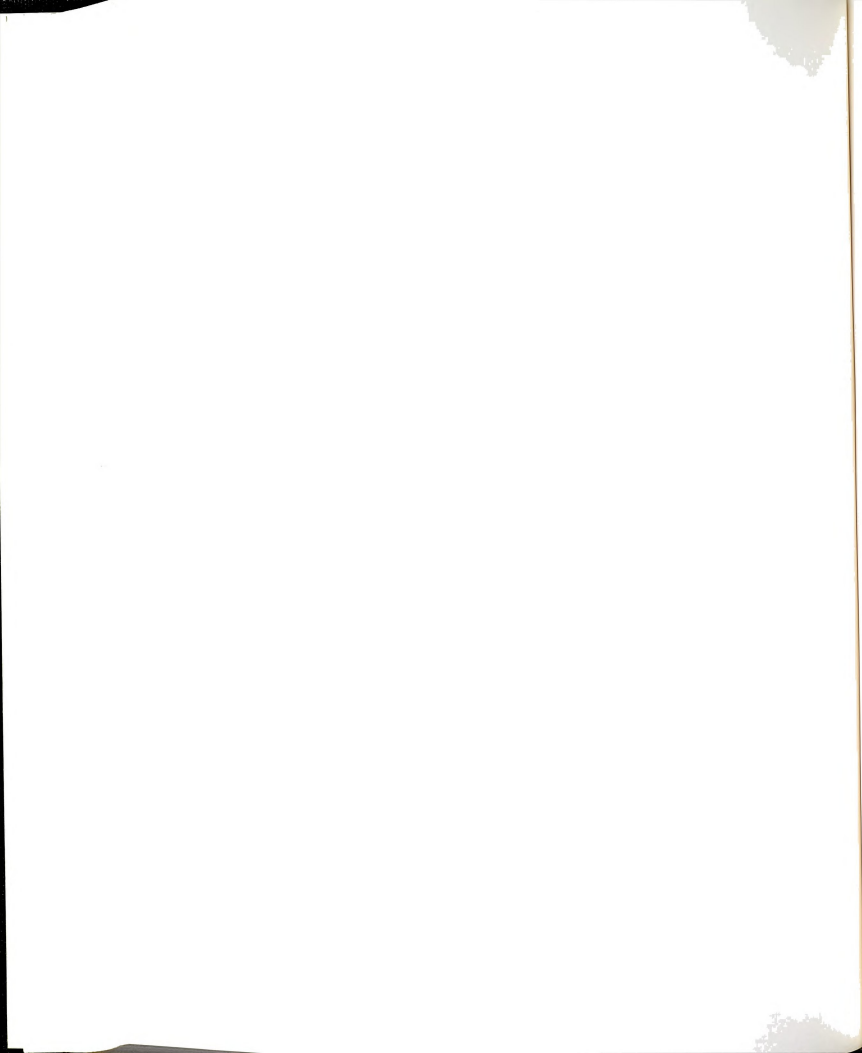
Clinical judgment can make use of the clinician's experience and knowledge in understanding an individual when some crucial information is lacking. A clinician can also decide the weights of the evidence. In a common statistical prediction framework, neither additional knowledge nor weights of the information can be utilized.

Hence, despite the large amount of evidence showing the superiority of statistical judgment (Kleinmuntz 1990), clinical judgment still has some important advantages over statistical prediction. Although people preferring statistical prediction can argue that artificial intelligence can replace all the above advantages of the intuitive elements in clinical judgment, yet no AI program can be developed in a foreseeable future that can replace all the judgmental functions of a physician or a clinical psychologist. Clinical judgment should also incorporate the knowledge of statistical prediction as part of the evidence for judgment. In this way, both the advantages of clinical and statistical judgement can be utilized. Research in making clinical judgment more explicit can help improve the mechanisms as well as the accuracy of clinical judgment. The errors of clinical judgment can be reduced also.

Cognitive Complexity as a Determinant of Judgment: A Hypothesis

The present study has revealed several situational factors which affect the mode of judgment people used, namely: diagnosticity, extreme base rate and physical mechanistic contexts. However, all of the personal variables including sex, age, major area and former knowledge about the Bayes' rule did not show any significant relationship with the mode of judgment used. Demographically, it seems that there is no effective way to predict whether a person is predominantly probabilistic or intuitive in their judgment.

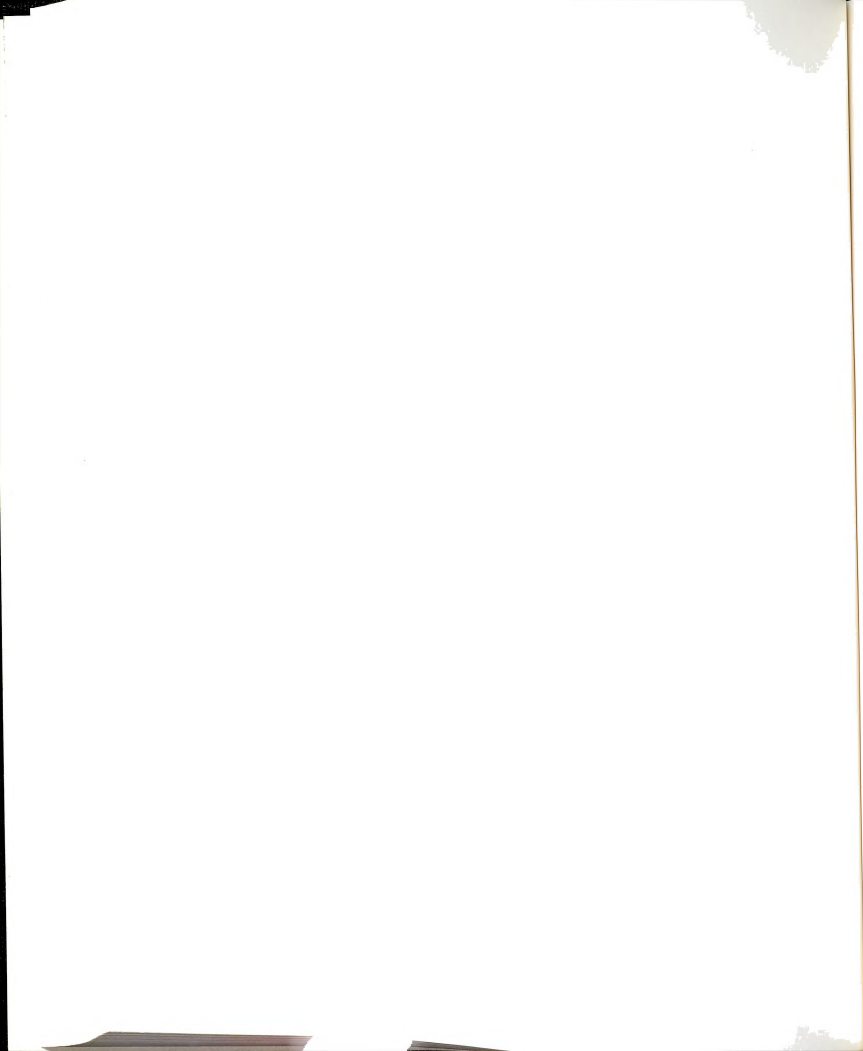
Nevertheless, some of our results showed some light of hint. The probabilistic subjects seemed to adopt a single strategy of judgment while the intuitive subjects changed more often. Qualitative analysis also showed that intuitive subjects attended to more differentiated aspects of the problems. The probabilistic subjects largely employed some over-simplified probabilistic rule on a limited amount of given information across different problems. Hence one possible personality determinant of judgment could be the cognitive complexity-simplicity of a person. Bieri et al. (1966, p.185) described: 'Cognitive complexity may be defined as the capacity to construe social behavior in a



multidimensional way. A more cognitive complex person has available a more differentiated system of dimensions for perceiving others' behavior than does a less cognitive complex individual'. As elaborated in the section on the qualitative dimensions of intuitive judgment, intuitive judgment seems to involve more complex evaluation and attends to a larger differentiated dimensions than probabilistic judgment does. The similarity of intuitive judgment and cognitive complexity on the differentiation of dimensions of judgment suggested that cognitive complex people might be more intuitive in their judgment under uncertainty than cognitive less complex people. Future research measuring the association of cognitive complexity with the mode of judgment might establish some useful personality determinants of the judgmental mode.

Implications for Future Research

Continuing research on intuitive judgment can be helpful in understanding human decision making as well as helping the design of machine intelligence. Qualitative studies on the decision making of experts in scientific and social affairs is recommended. Attention can be paid to how people make errors in using the cognitive and meta-cognitive skills of intuitive judgment, but not how people make errors



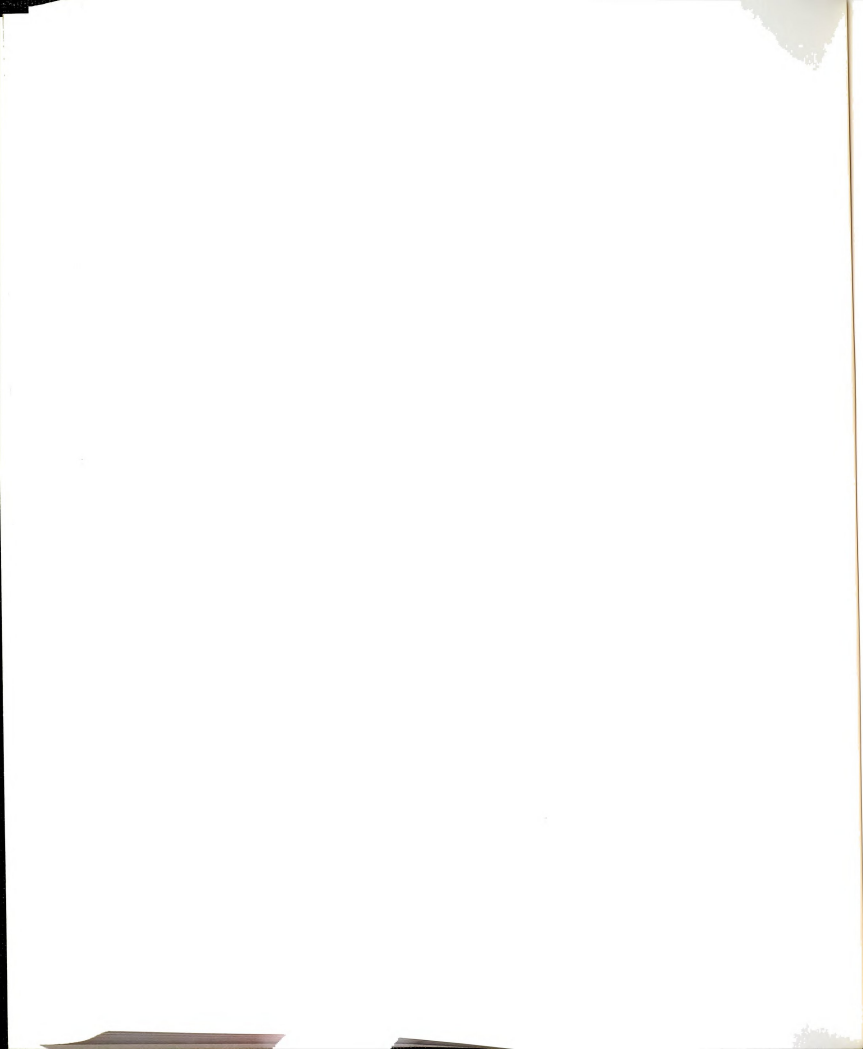
against the mathematical and logical norms within a theoretical world.

Different decision situations, besides the base rate problem, such as employment, marriage and business decision making can be studied by creating scenarios for people to judge and by recording the reasonings of the people in making judgment.

The methods of establishing constructs through qualitative methods and the subsequent application of quantitative analysis is recommended to future research in the field of psychology of judgment, as well as for other fields in the social sciences.

As hypothesized, cognitive complexity might be related to probabilistic judgment as a personality determinant. Future research should be directed to test this hypothesis.

Since the subjects in this study were a special group of overseas Chinese graduate students, generalizations of the results here to the whole Chinese or American population might not be immediate. Similar research using other local Chinese subjects as well as American subjects should be studied and compared with the present results.

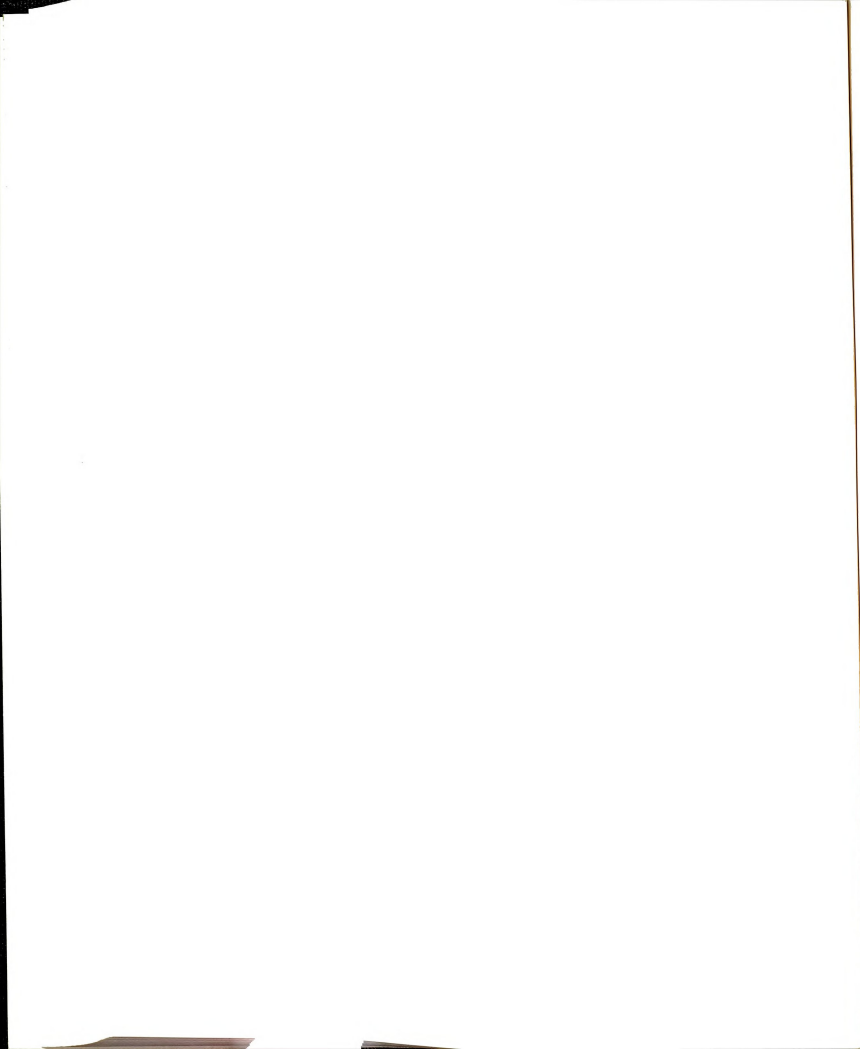


Limitations

The present sample of subjects consisted of mainly Chinese overseas graduate students in Michigan State University. Although the numerical responses between our sample and those of the past research using Western subjects are highly similar, the hasty generalization to the similar American population is not advisable until confirmation by collection of similar verbal reasoning data on the American subjects.

Despite the highly reliable property of the coding scheme, the usefulness of the qualitative categories of the two modes of judgment need to be ascertained by future research in the similar area.

The present experimental design, as demonstrated, could reveal the properties of intuitive judgment which can make a case for the defend of intuitive rationality against the attacks by earlier researchers. However, whether the intuitive judges had rightfully used the base rate information has to be investigated within the framework of intuitive judgment, not from a faultless and mechanical normative environment. Therefore whether there is truly a defect in intuitive judgment in making use of the base rate could not be empirically tested until the mode of intuitive judgment is adequately described by further research along some similar lines of the present research.



Appendix A

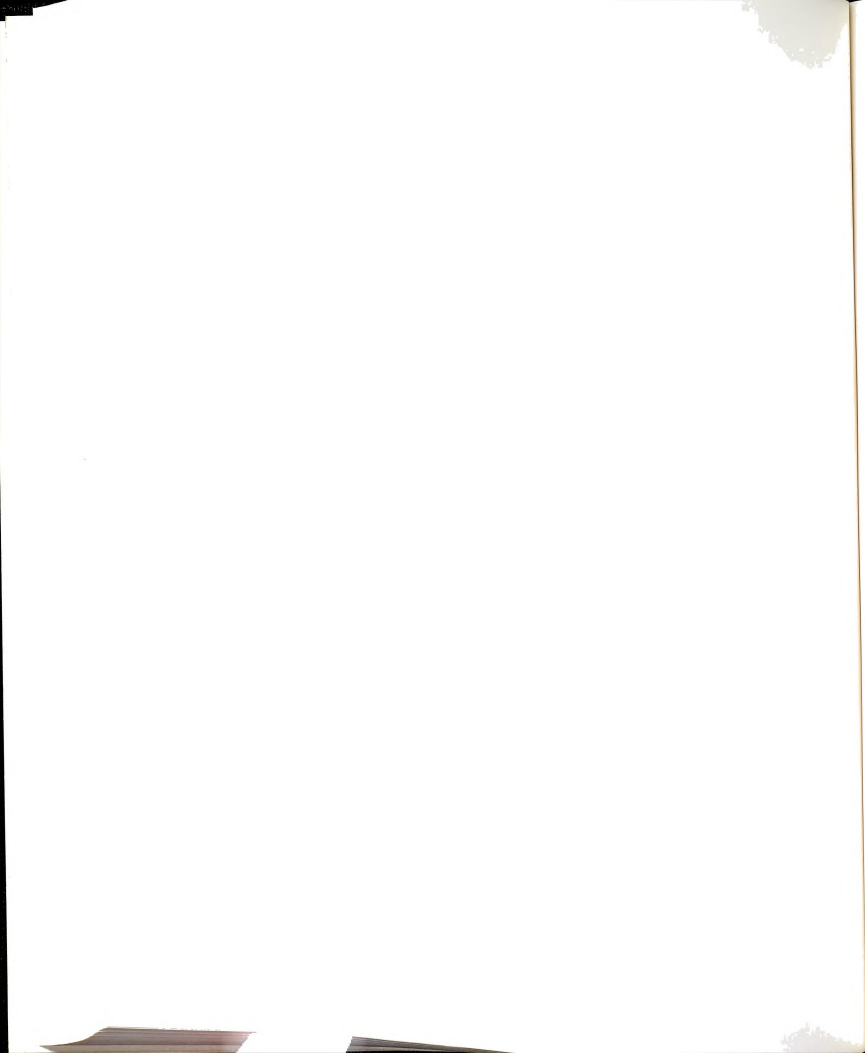
Questionnaire of the Base Rate Problems

Instructions

"This study is to investigate the thinking process of people's judgement of some uncertain everyday affairs. Please literally read the questions and try to think aloud. I am not primarily interested in your final solution, still less in your reaction time, but in your thinking behavior, in all your attempts, in whatever comes into your mind, to recount exactly what unfolds in your consciousness, your hesitations, doubts, the ideas which come into your mind, etc. Be bold and speak them out."

"I can assure you that the study has nothing to do with the study of your IQ or personality, unlike many psychological experiments, there is no deception of any kind. There is no so-called correct answer to the questions. Your response would not be judged as right or wrong."

"While you are thinking aloud or answering during the interview, the content will be tape-recorded for detailed analysis. If you have any questions or any words you don't understand, feel free to ask me. ... Do you have any questions now? If not, or shall we begin?"



(A)

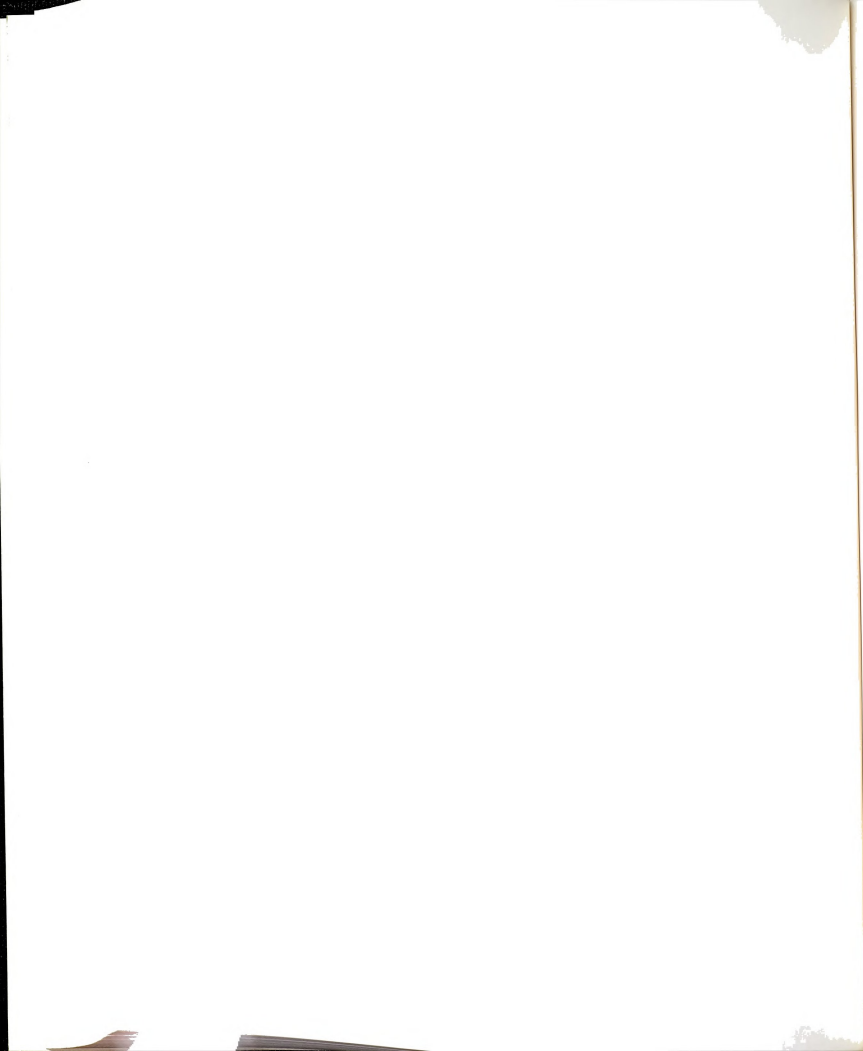
"A car was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

(a) 85% of the cabs in the city are Green and 15% are Blue.

(b) a witness identified the Cab as Blue. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time.

What is the probability that the cab involved in the accident was Blue rather than Green?"

Answer: _____%



(B)

This problem is the same as the last one in all aspects except that sentence (a) has been modified.

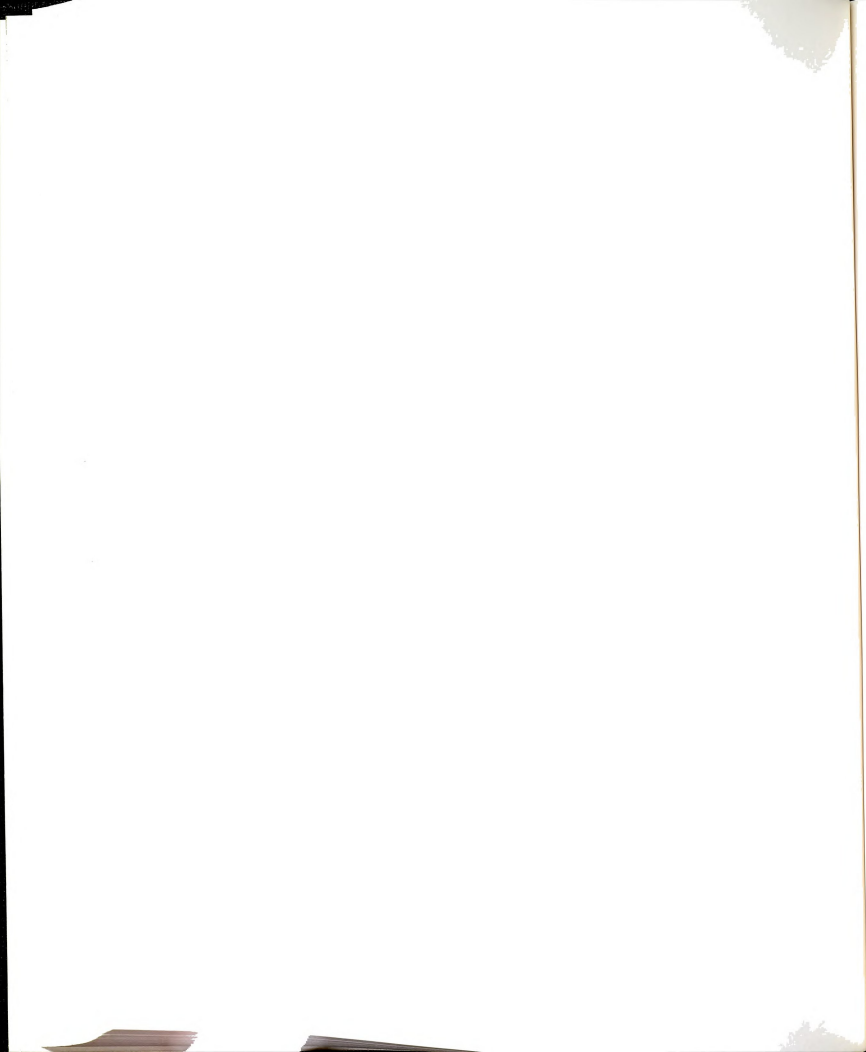
"A car was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

(a') Although the two companies are roughly equal in size, 85% of cab accidents in the city involve Green cabs and 15% involve Blue cabs."

(b) a witness identified the Cab as Blue. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time.

What is the probability that the cab involved in the accident was Blue rather than Green?"

Answer: _____%



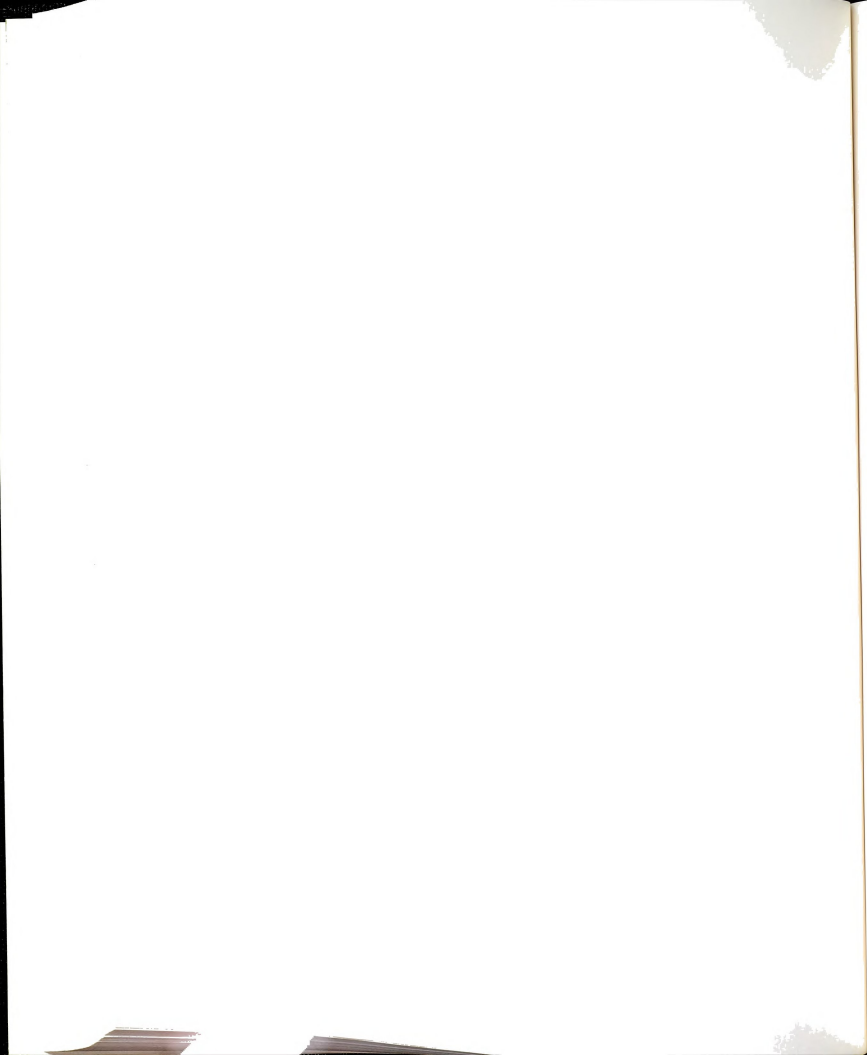
(8)

A panel of psychologists have interviewed and administered personality tests to 30 engineers and 70 lawyers, all successful in their respective fields. On the basis of this information, thumbnail descriptions of the 30 engineers and 70 lawyers have been written. You will find below a description, chosen at random from the 100 available descriptions.

"Jack is a 45-year-old man. He is married and has four children. He is generally conservative, careful, and ambitious. He shows no interest in political and social issues and spends most of his free time on his many hobbies which included home carpentry, sailing, and mathematical puzzles."

The probability that Jack is one of the 30 engineers in the sample of 100 is :

Answer: _____%



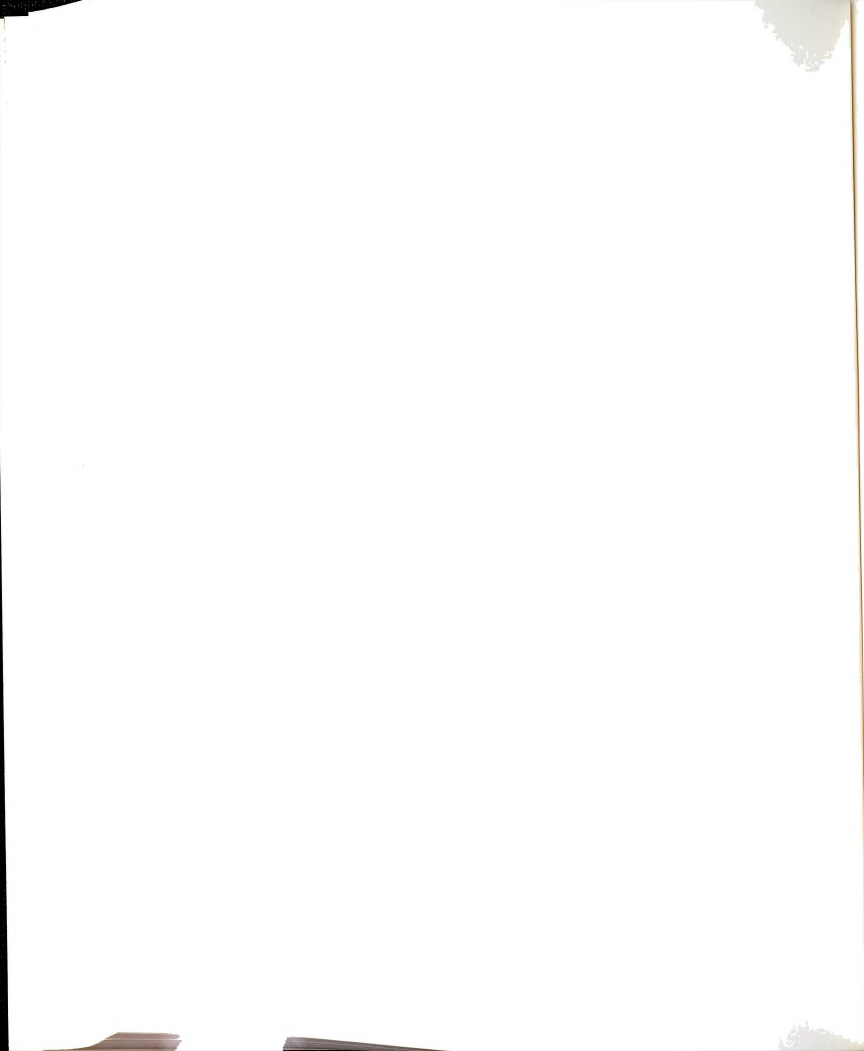
(D)

Everything being the same as the last problem, please consider another description drawn from the same group of people:-

"Dick is a 30-year-old man. He is married with no children. A man of high ability and high motivation, he promises to be quite successful in his field. He is well liked by his colleagues.

The probability that Dick is one of the 30 engineers in the sample of 100 is:

Answer: _____%

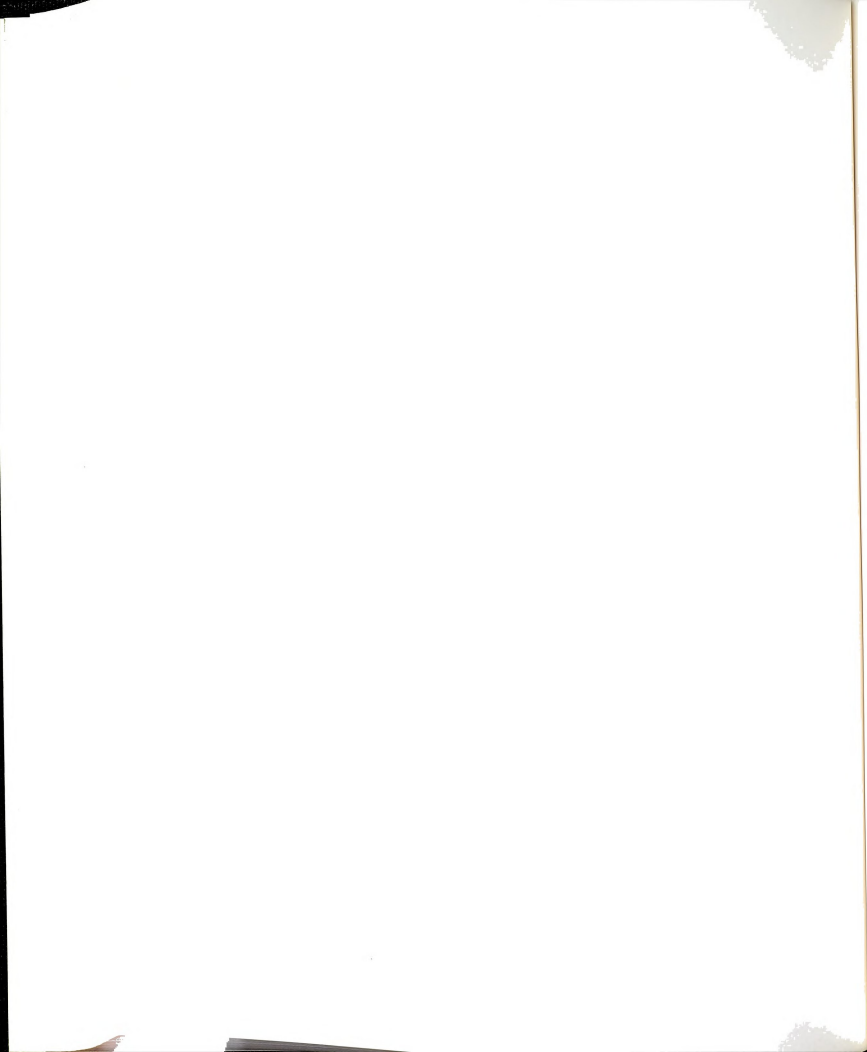


(E)

Please consider the problem (c) again, supposing that instead of 30, there is only one engineer in the group, all others being lawyers.

The probability that Jack is that only engineer in the sample of 100 is :

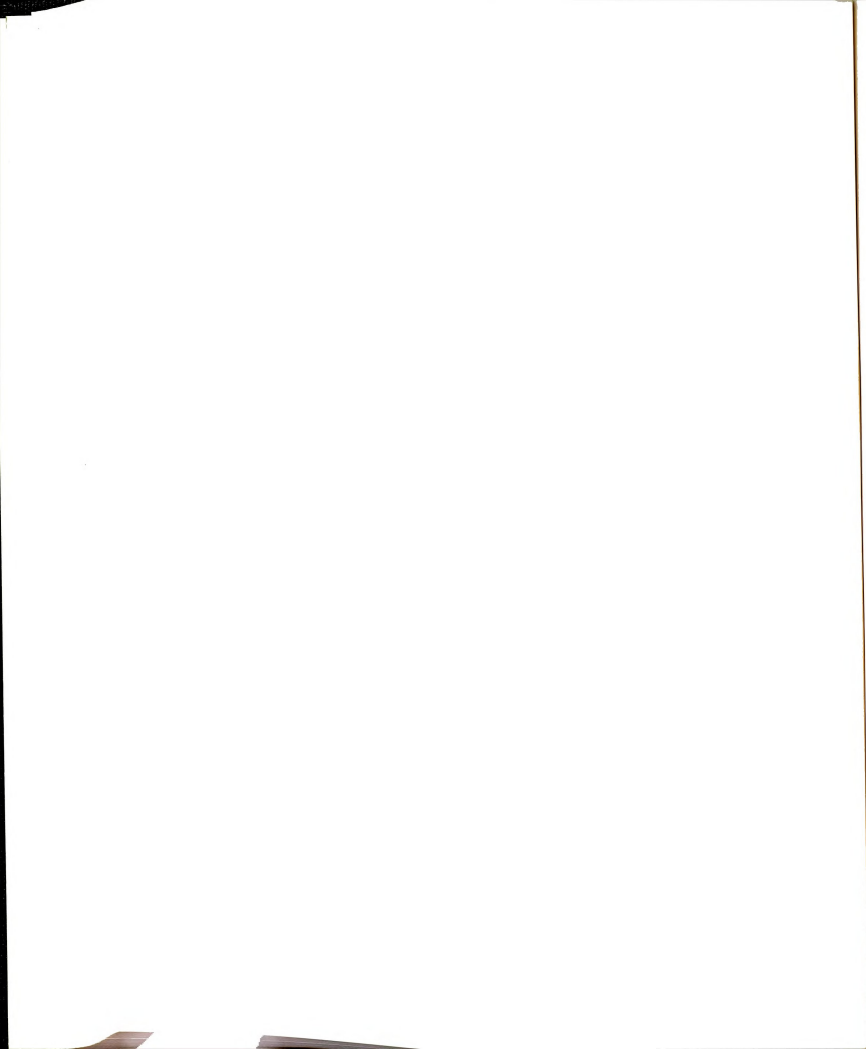
Answer: _____%



(F)

A still developing machine of computer vision will commit error randomly about 30% of any time. Suppose the machine declares a certain figure to be an ellipse while given a trial document containing 90 circles and 10 ellipses, estimate the probability that the figure is really an ellipse.

Answer: _____%

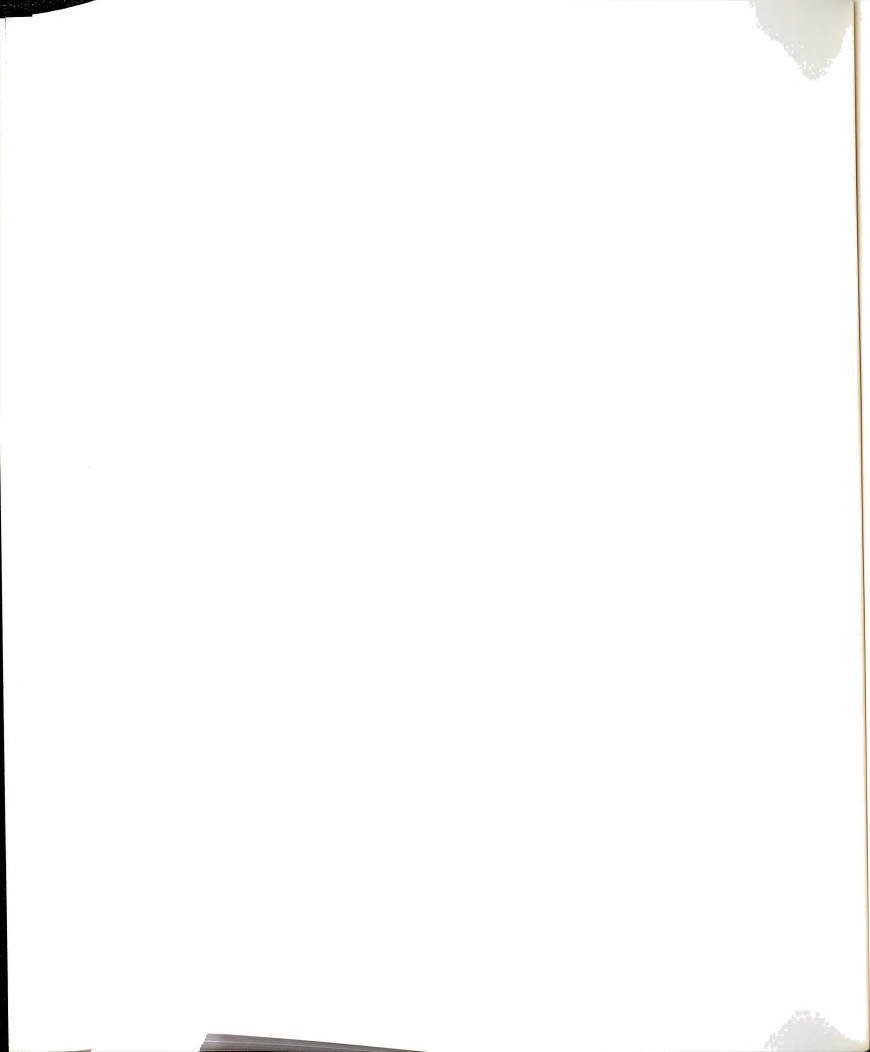


(6)

Given the information that (i.e. knowing that) the event B has occurred, the probability for the occurrence of the event A is $6/10$. Knowing that B does not occur, the probability for the occurrence of the event A is 0.4 . The natural probability of occurrence for the event B is $1/10$.

Now given that (i.e. knowing that) the event A has occurred, what is the probability for the occurrence of event B?"

Answer: _____%



APPENDIX B
Two Protocol Examples

INTERVIEW PROTOCOL

Subject No: 38

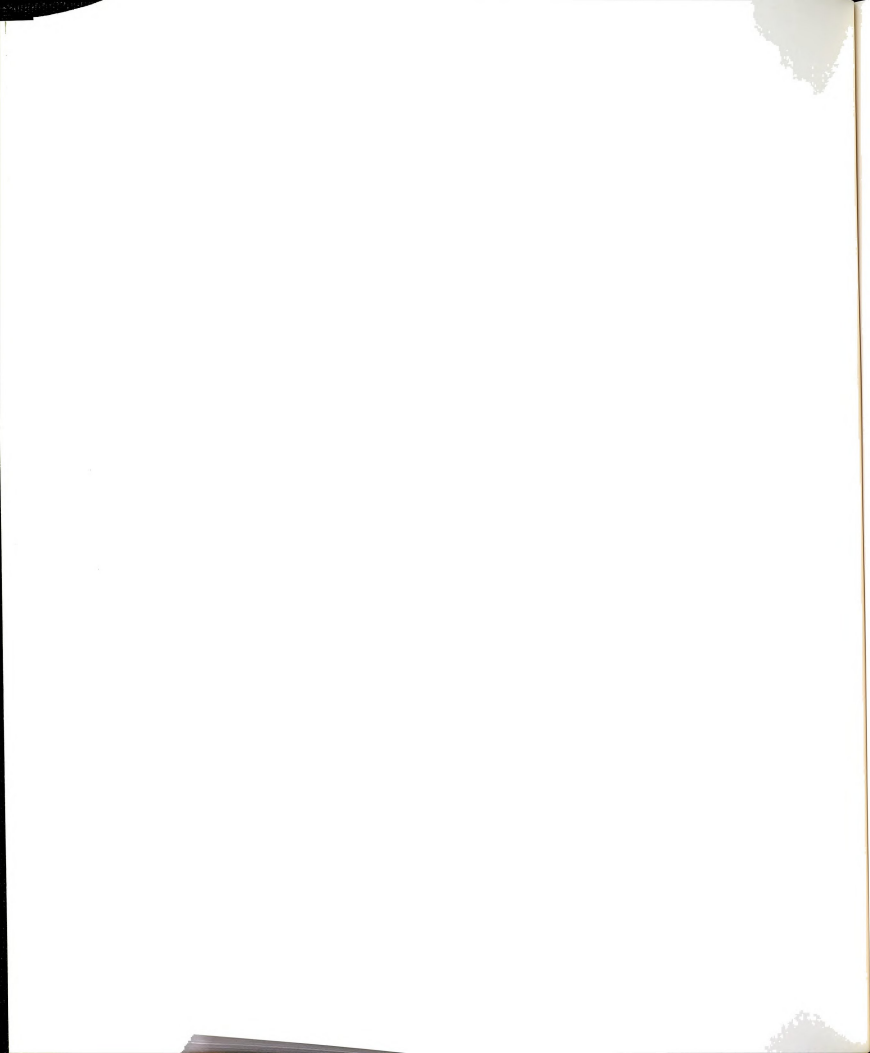
Sex: M

A) There are two cases involved: green and failure; blue and success. The probability is $85\% \cdot 20\% + 15\% \cdot 80\% = 29\%$.

B) Method same as in (A).

C) The only clue is that he shows no interest in political and social issues and ... like mathematical puzzles. He seems more like an engineer. It implies that the probability is $30/100$. And since he is chosen from 100 people. The probability is $1\% \cdot 30\% = 3\%$.

D) The clue about him is even fewer, it is hard to judge. It is hard to use calculations. Thinking with numbers here has little use. The best way is not to calculate. Given information is too few. <No answer, the subject does not want to give an answer to this problem>



E) $1/100 * 1/100 = 0.01\%$. Picking one from 100, its own probability is $1/100$. (reasoning is similar to that in problem C) Therefore the probability is 0.01% .

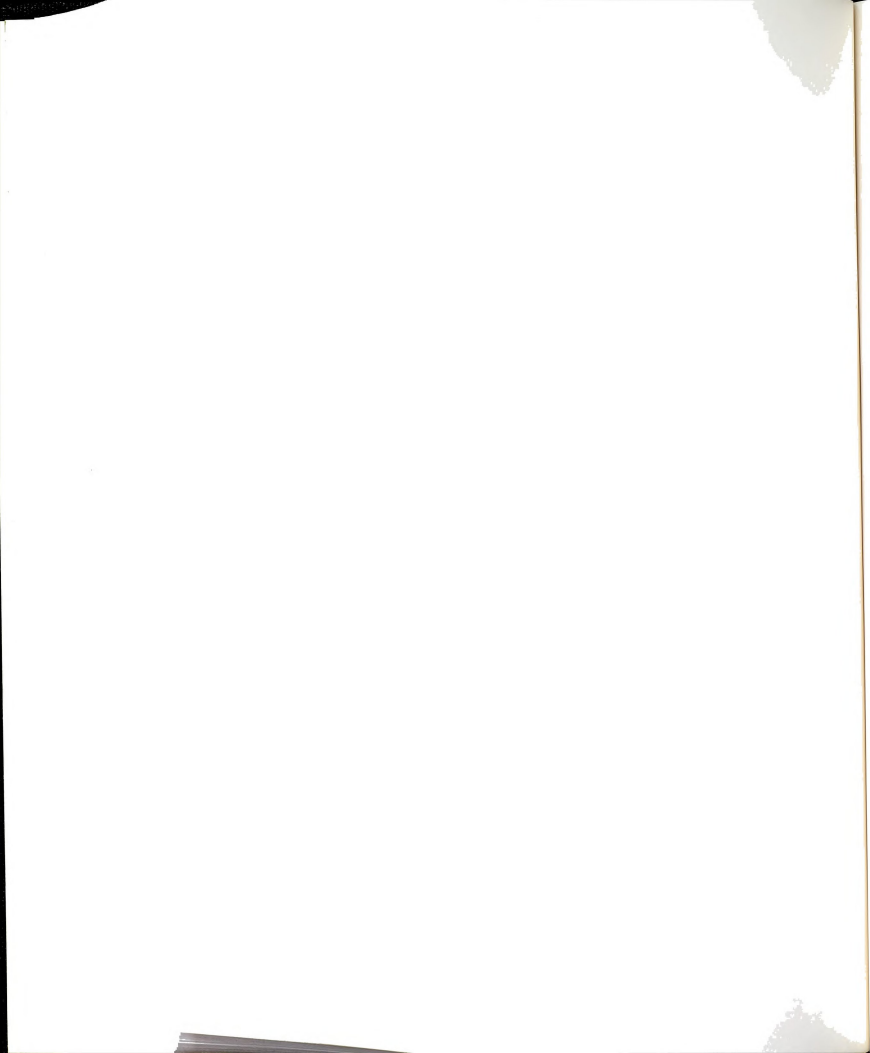
F) $70/100 * 10/100 = 7\%$.

G) $P(AB)=0.6$, $P(AB')=0.4$, $P(B)=1/10$. $P(B/A)=P(AB)/P(A)=0.4*0.01/(0.4*0.1+0.6*0.9)=6.8\%$. I haven't used Bayes' theorem for several years already.

Statistical knowledge: One course of statistics at 2nd year at undergraduate college.

Knowledge about Bayes' Theorem: learned before, remember somewhat.

Major: Computer Science.



INTERVIEW PROTOCOL

Subject No: 40

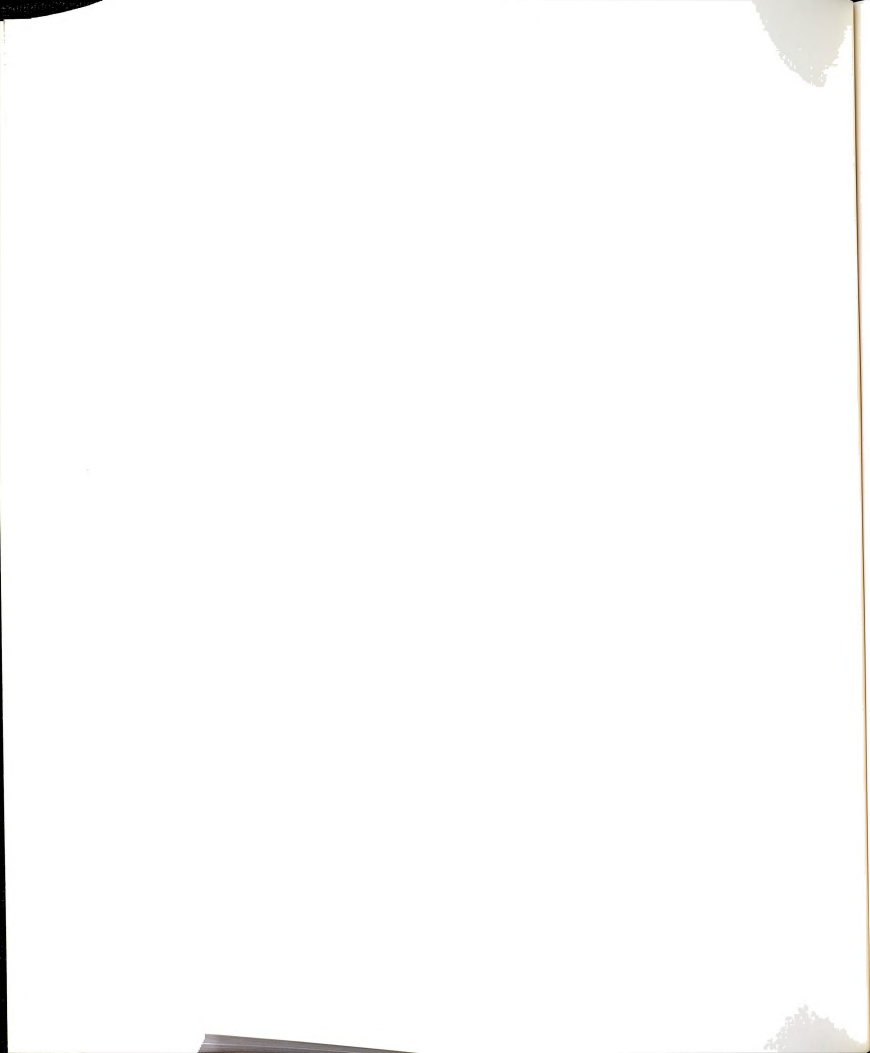
Sex: F

A) I believe in the witness, believe in what he said. The degree of belief is 80%. The occurrence rate of 85:15 has little influence. We should believe in the witness. < 80% >

B) Here I am more certain. Accident rate is similar to the past crime record. It is not right to suspect him (the one with crime record) when we have a crime incident. Still 80%.

c) Shows no interest in political and social issues. A Lawyer should care about the society, care about politics. Originally it should be 50:50. Since he shows no interest, it becomes 25:75; he likes mathematical puzzles, therefore it becomes 12: 87. (taking a further half from 25) The probability is 87%.

D) He works well with his colleagues. Lawyers always work together with colleagues, share and exchange opinions. Engineers work more by himself. It is possible to be a lawyer. The answer is 13% (using a similar method as problem C, $100\% - 87\% = 13\%$). It only means the probability is very small, the numbers doesn't mean very much. It only means it is very small.



E) This is the same as doing the accident rate of cabs. It (the ratio) doesn't have any meaning. The probability is 87%. I trust my judgment.

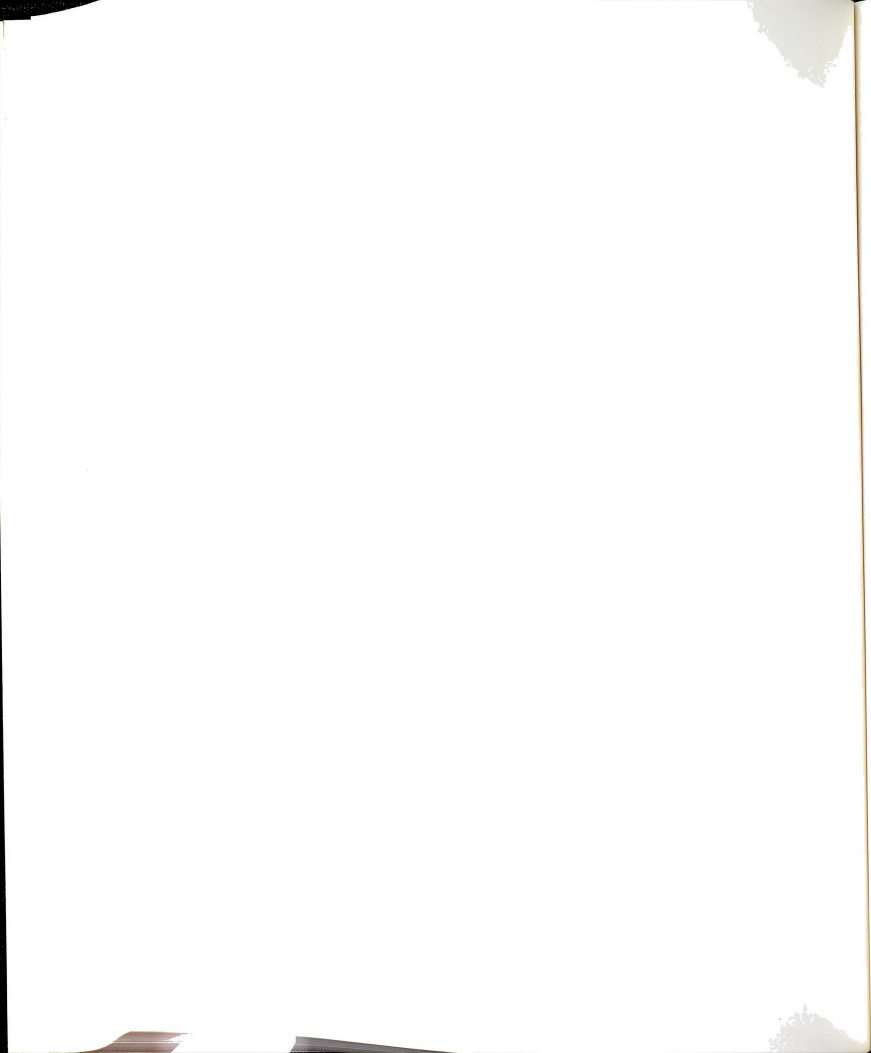
F) I have little concepts about numbers. Since the error rate is 30%, the hit rate should be 70%. The method of reasoning is the same as before (same as problem A, B). < 70% >

G) 40%. B occurs and A occurs is 60%. Therefore A occurs and B occurs is 40%.

Statistical knowledge: afraid of mathematics, learned nothing about statistics.

Knowledge about Bayes' Theorem: Never heard about it.

Major: Musicology.



APPENDIX C

Classification of the Two Protocols

Here is an interview protocol from subject #38, a male graduate student in computer science:

A) There are two cases involved: green and failure; blue and success. The probability is $85\%*20\%+15\%*80\% = 29\%$.

Classification: probabilistic

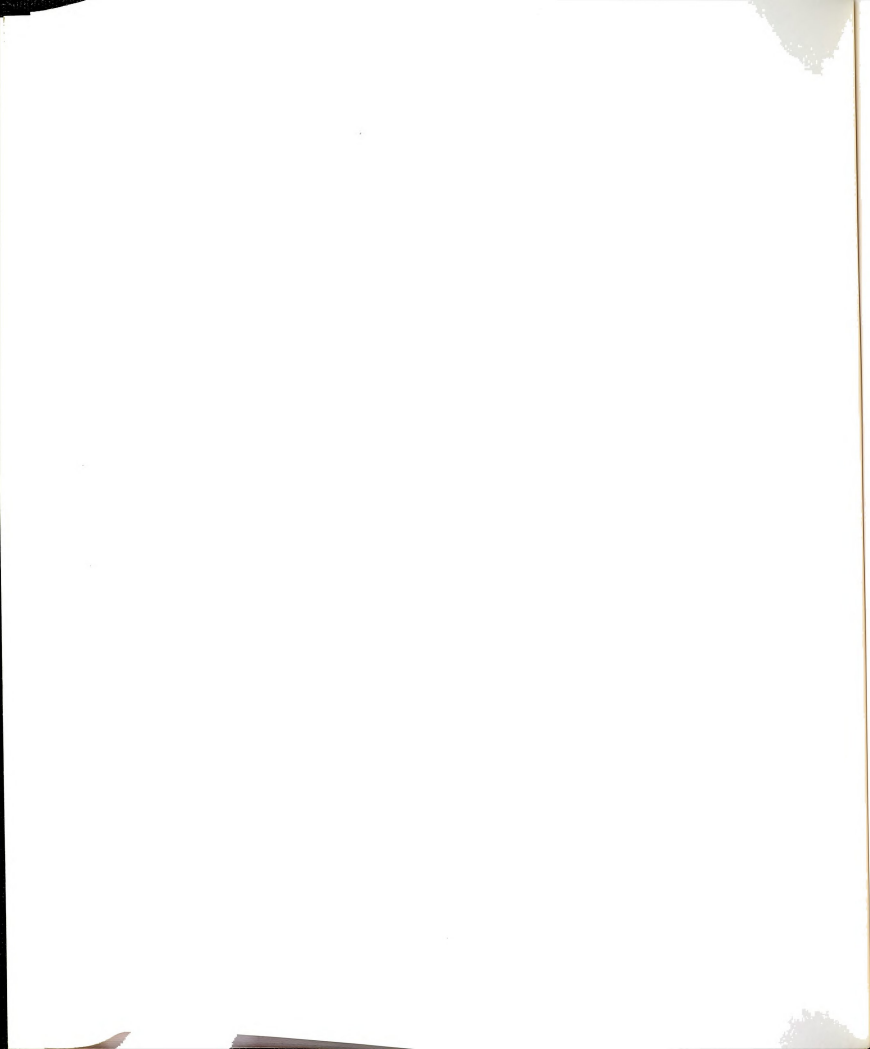
Reason: subject explicitly employs the additive and multiplicative rules of axiomatic probability

B) Method same as in (A).

Classification: probabilistic

Reason: same as above

C) The only clue is that he shows no interest in political and social issues and ... like mathematical puzzles. He seems more like an engineer. It implies that the probability is 30/100. And since he is chosen from 100 people. The probability is $1\%*30\%=3\%$.



Classification: probabilistic

Reason: subject actively employs the urn model (i.e. 30/100) for determining the probability of selecting an engineer in a group of 100 people, 30 of which are engineers. Besides, there is heavy reliance on the multiplicative rule of probability theory.

D) The clue about him is even fewer, it is hard to judge. It is hard to use calculations. Thinking with numbers here has little use. The best way is not to calculate. Given information is too few. <No answer, the subject does not want to give an answer to this problem>

Classification: missing

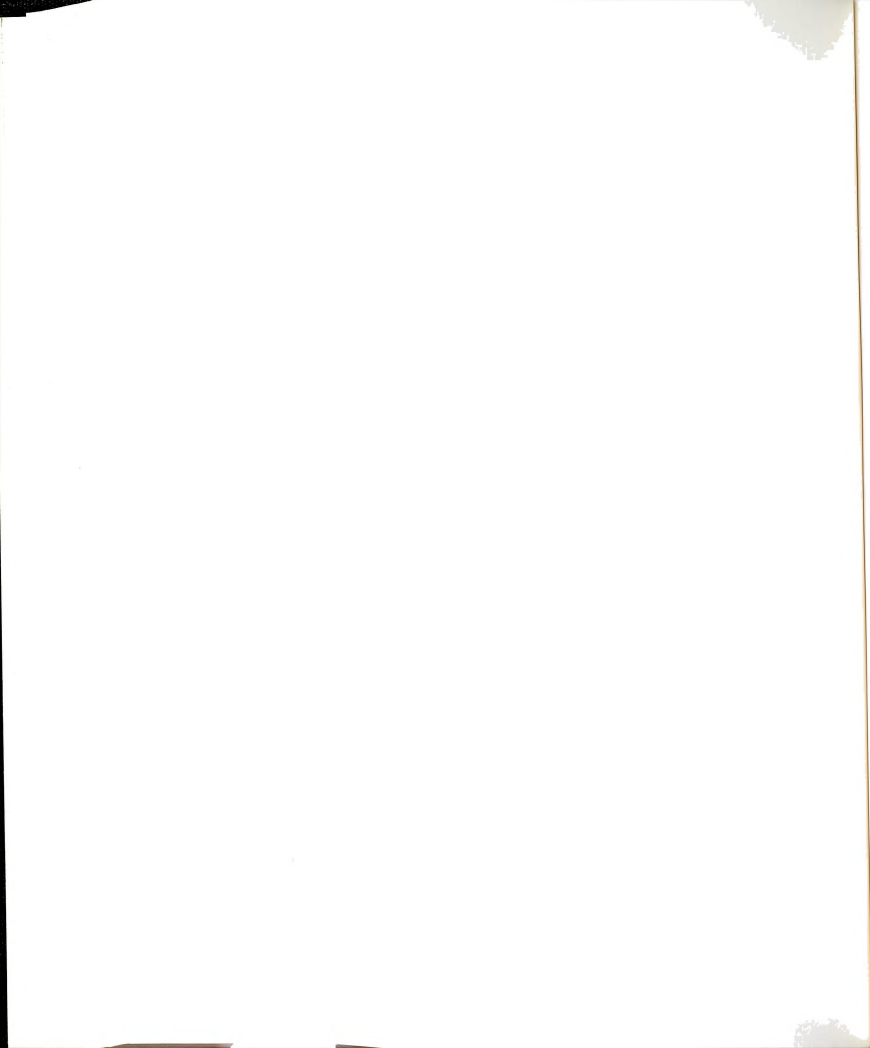
Reason: subject decides that it is appropriate for him to give any answer to this problem.

E) $1/100 * 1/100 = 0.01\%$. Picking one from 100, its own probability is 1/100. (reasoning is similar to that in problem C) Therefore the probability is 0.01%.

Classification: probabilistic

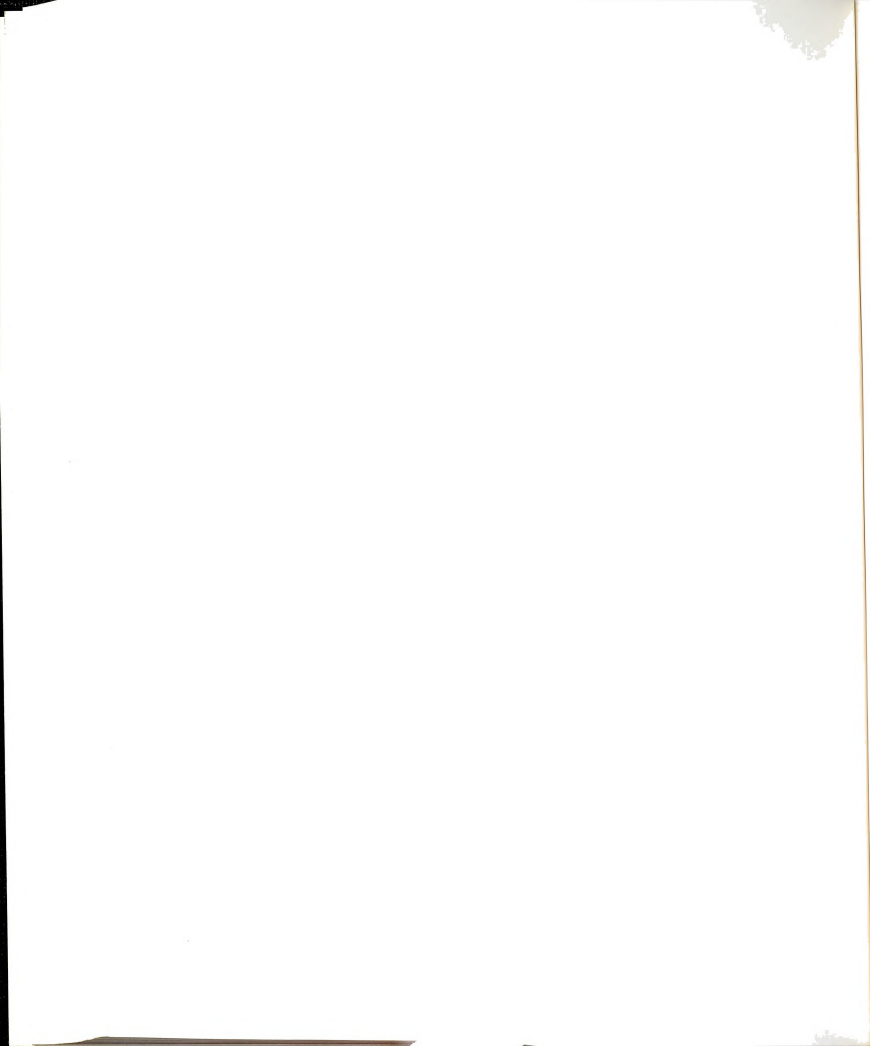
Reason: subject's pattern of reasoning is the same as in problem C above.

F) $70/100 * 10/100 = 7\%$.



Classification: probabilistic

Reason: subject solely relies on axiomatic probability theories.



And the following protocol comes from a female student (subject # 40) with a major in music:

A) I believe in the witness, believe in what he said. The degree of belief is 80%. The occurrence rate of 85:15 has little influence. We should believe in the witness. < 80% >

Classification: intuitive

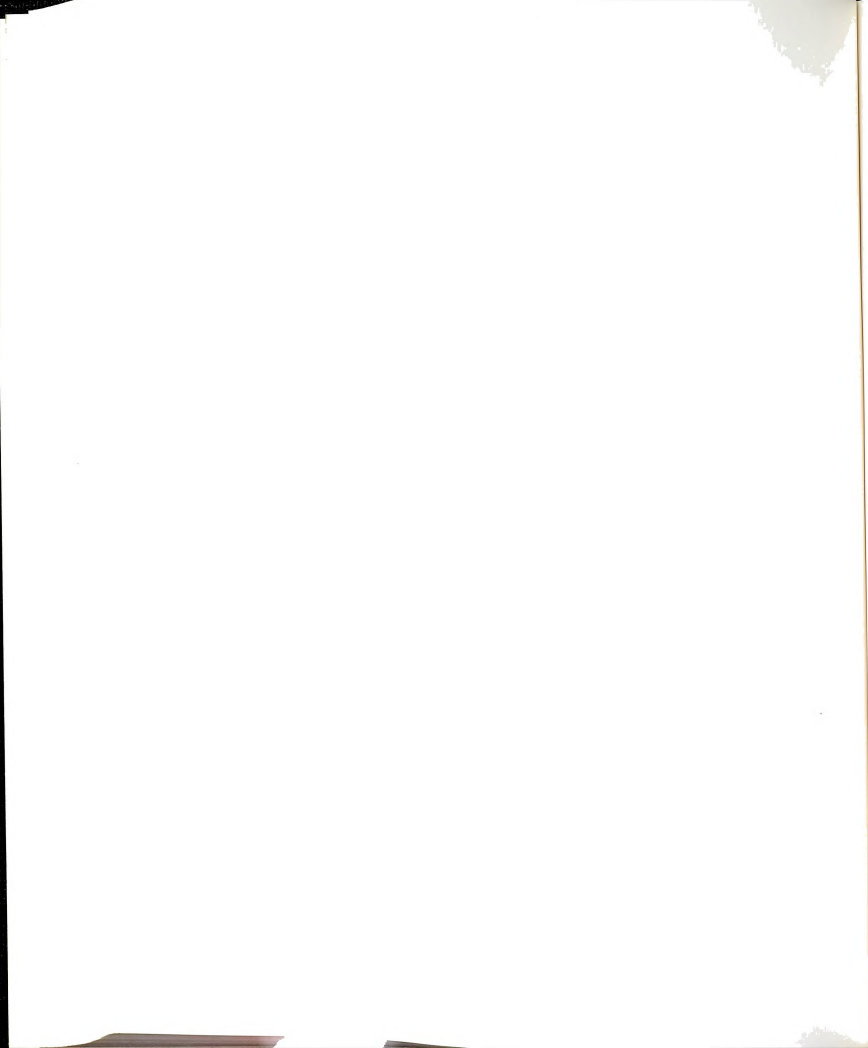
Reason: subject does not employ any probability theory, consider the base rate as having little or no influence, trust the individual case (witness) than the statistical data.

B) Here I am more certain. Accident rate is similar to the past crime record. It is not right to suspect him (the one with crime record) when we have a crime incident. Still 80%.

Classification: intuitive

Reason: subject again does not rely on probability theory, consider the past as not determining the present case, believe in the individual witness information.

c) Shows no interest in political and social issues. A Lawyer should care about the society, care about politics. Originally it should be 50:50. Since he shows no interest, it becomes 25:75; he likes mathematical puzzles, therefore



it becomes 12: 87. (taking a further half from 25) The probability is 87%.

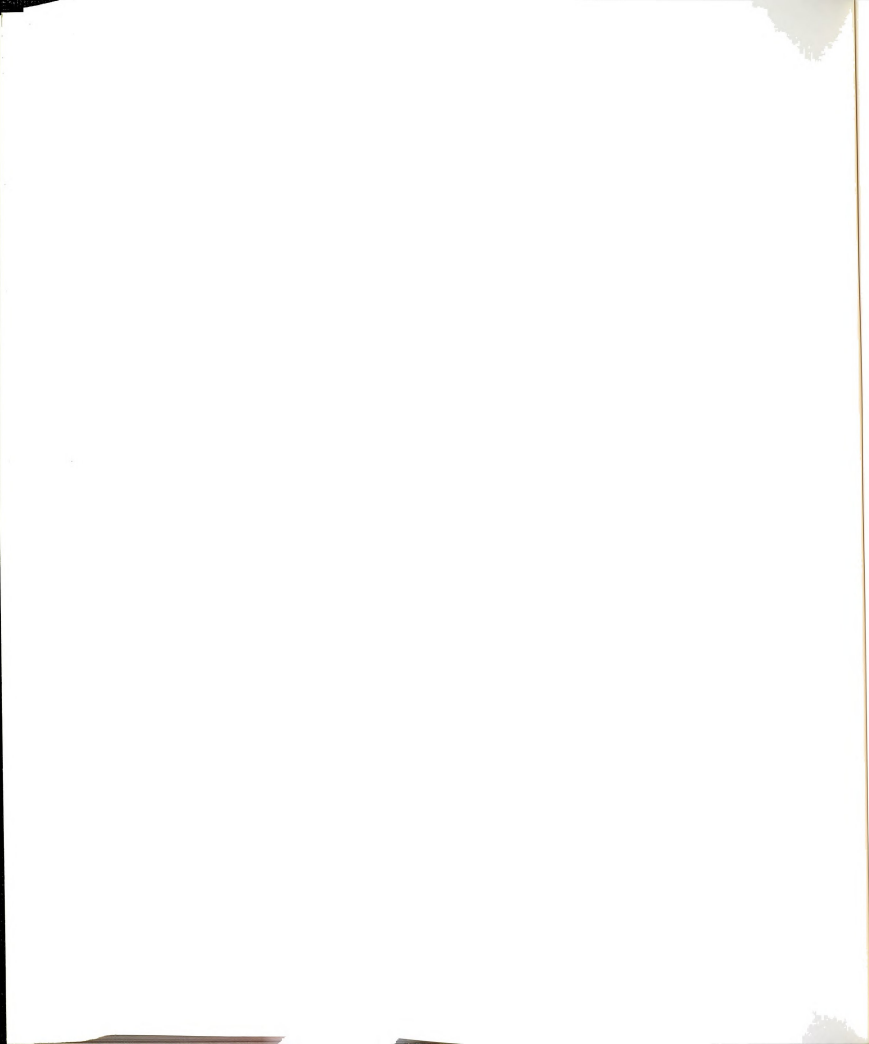
Classification: intuitive

Reason: subject does not rely on any formal probability theory. Although subject does manipulate with some numbers, she does it in a self-made way so as to express her degree of belief with respect to the evidence. She also uses the IF-THEN criterion : if one is a lawyer, one should care about politics, to determine the career from the given information.

D) He works well with his colleagues. Lawyers always work together with colleagues, share and exchange opinions. Engineers work more by himself. It is possible to be a lawyer. The answer is 13% (using a similar method as problem C, $100\% - 87\% = 13\%$). It only means the probability is very small, the numbers doesn't mean very much. It only means it is very small.

Classification: intuitive

Reason: creating narratives of a normal working atmosphere of a lawyer and engineer to fill in the gaps of information in this case. Subject does not predominantly rely on some probability theories.



E) This is the same as doing the accident rate of cabs. It (the ratio) doesn't have any meaning. The probability is 87%. I trust my judgment.

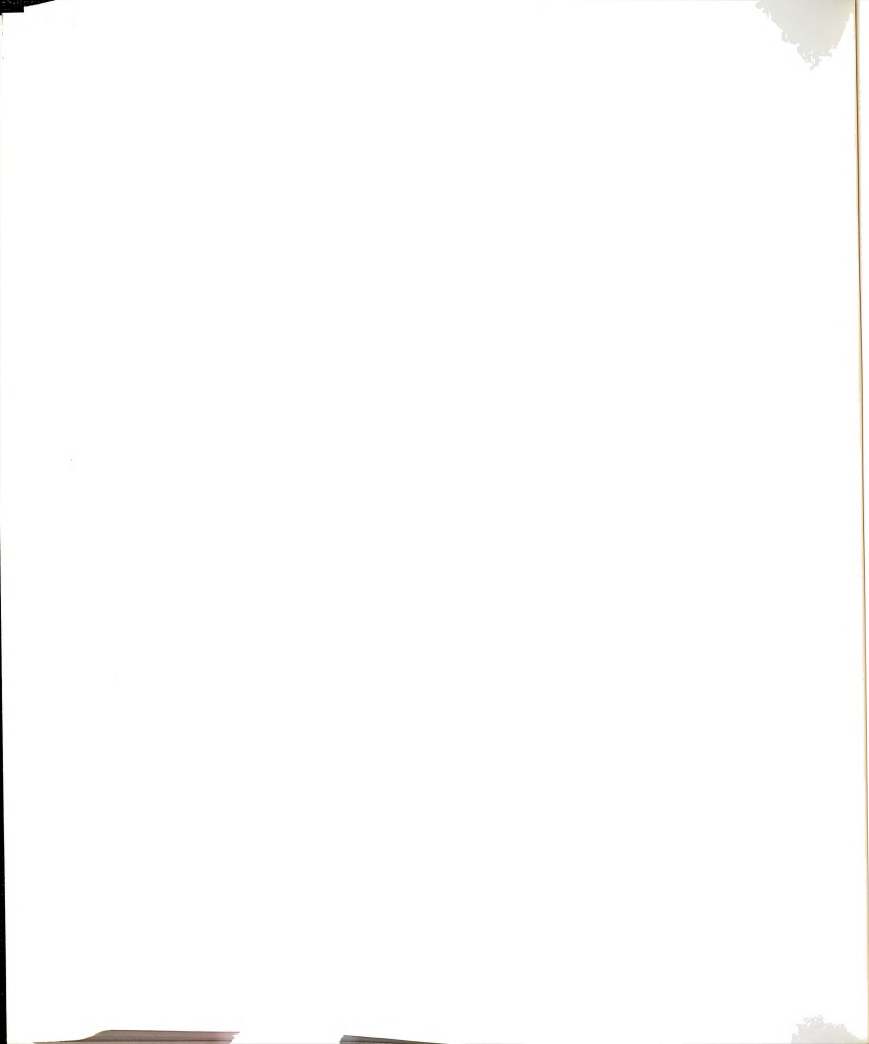
Classification: intuitive

Reason: reason is analogous to those for problem A or B above

F) I have little concepts about numbers. Since the error rate is 30%, the hit rate should be 70%. The method of reasoning is the same as before (same as problem A, B). $< 70\% >$

Classification: intuitive

Reason: reason similar to those in problem A or B. Subject relies heavily on information about the individual case concerned, and not rely on probability theory. Surely the subject has to know at least that hit rate = $100\% - \text{error rate}$, in order to comprehend the question. The possession of such knowledge does not mean that the judgment is predominantly a probabilistic mode.



APPENDIX D

Histograms of the Numerical Judgmental Responses

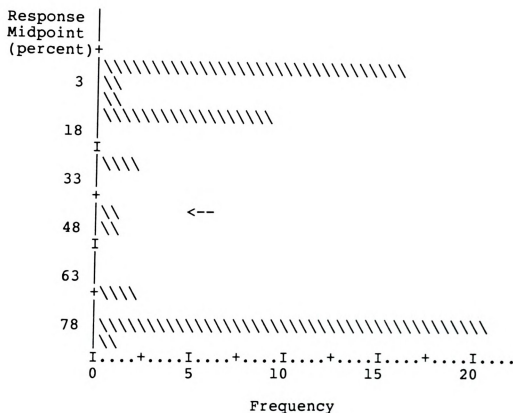
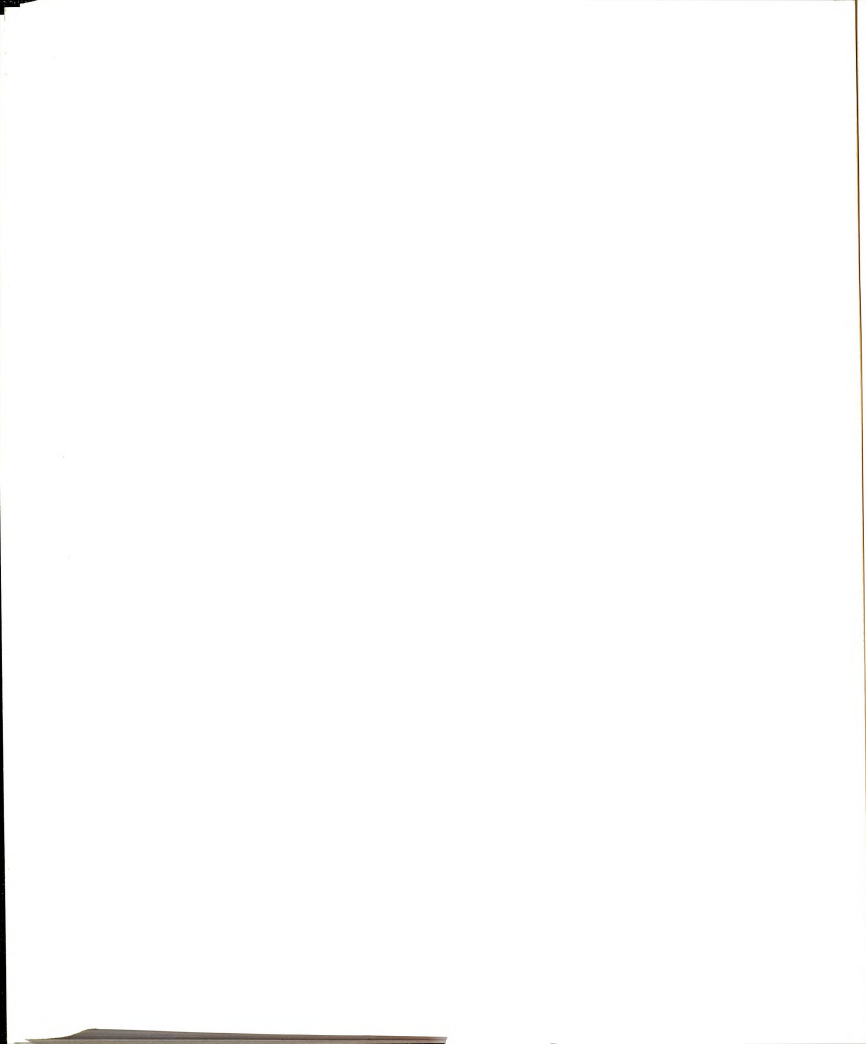


Figure D.1: Histogram of the subjects' responses for problem A. Bayes' estimate denoted by the arrow sign <--.



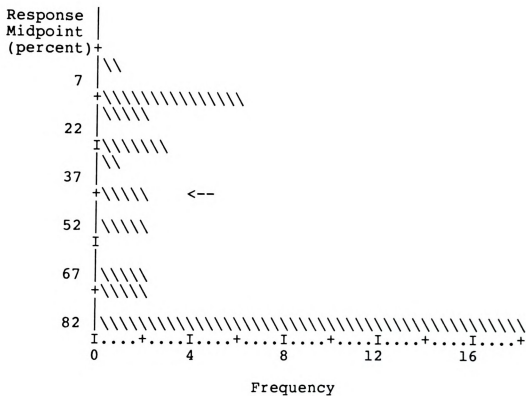
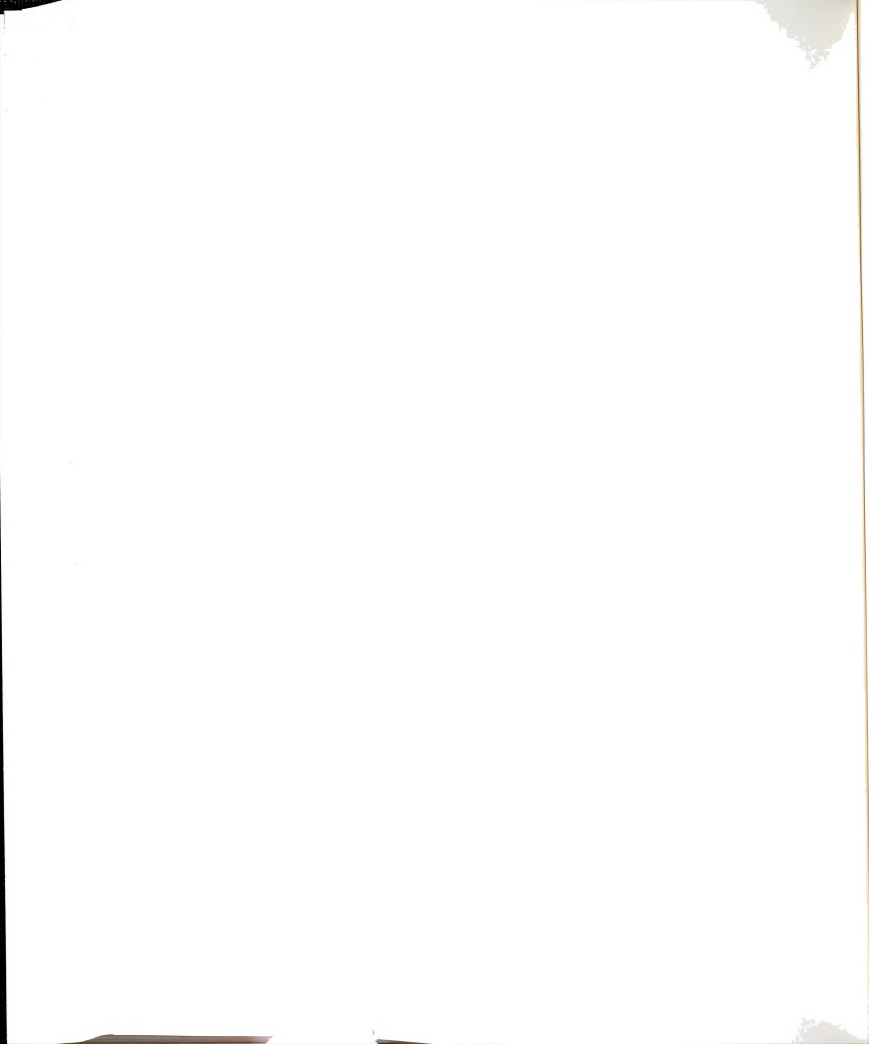


Figure D.2: Histogram of the subjects' responses for problem B. Bayes' estimate denoted by the arrow sign <--.



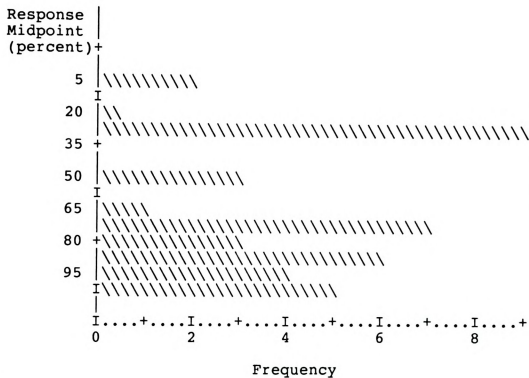


Figure D.3: Histogram of frequency of subject's response for problem C. Bayes' estimate not available for this problem.

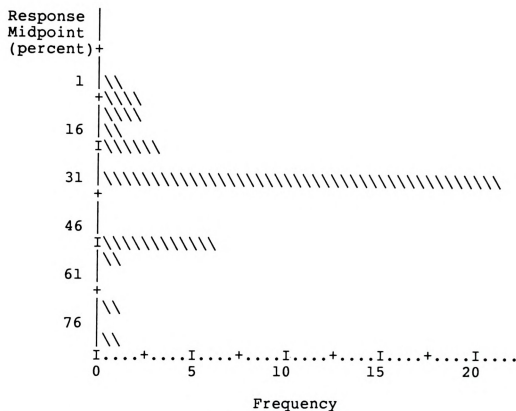


Figure D.4: Histogram of the subjects' responses for problem D. Bayes' estimate not available for this problem.

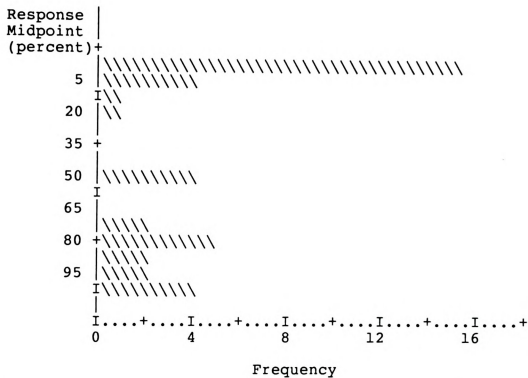
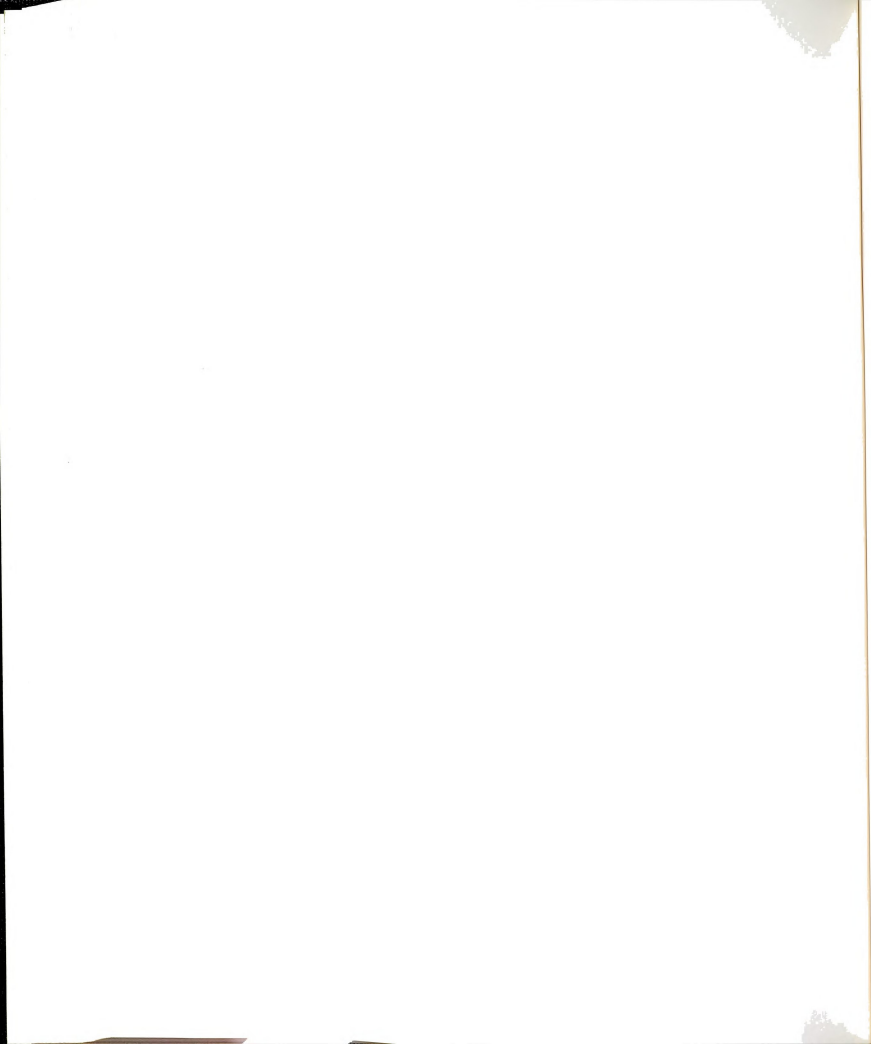


Figure D.5: Histogram of the subjects' responses for problem E. Bayes' estimate not available for this problem.



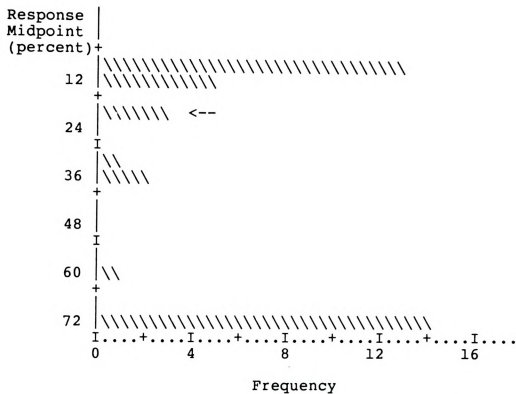
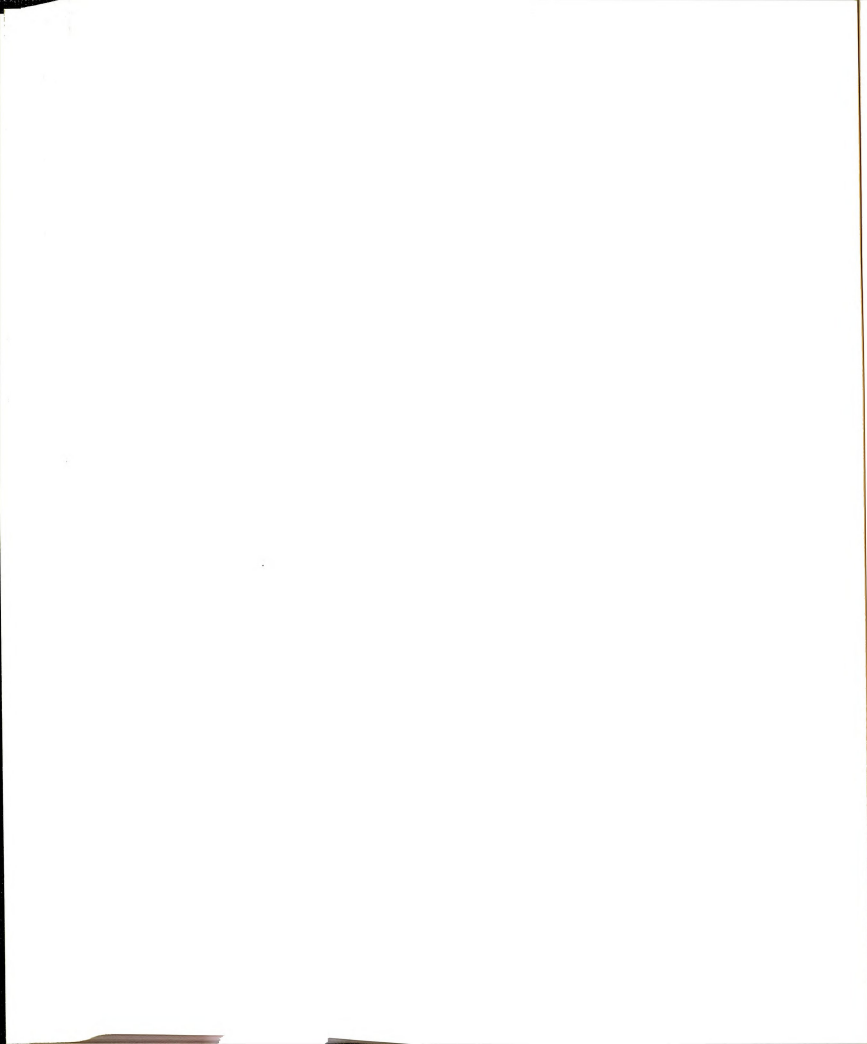


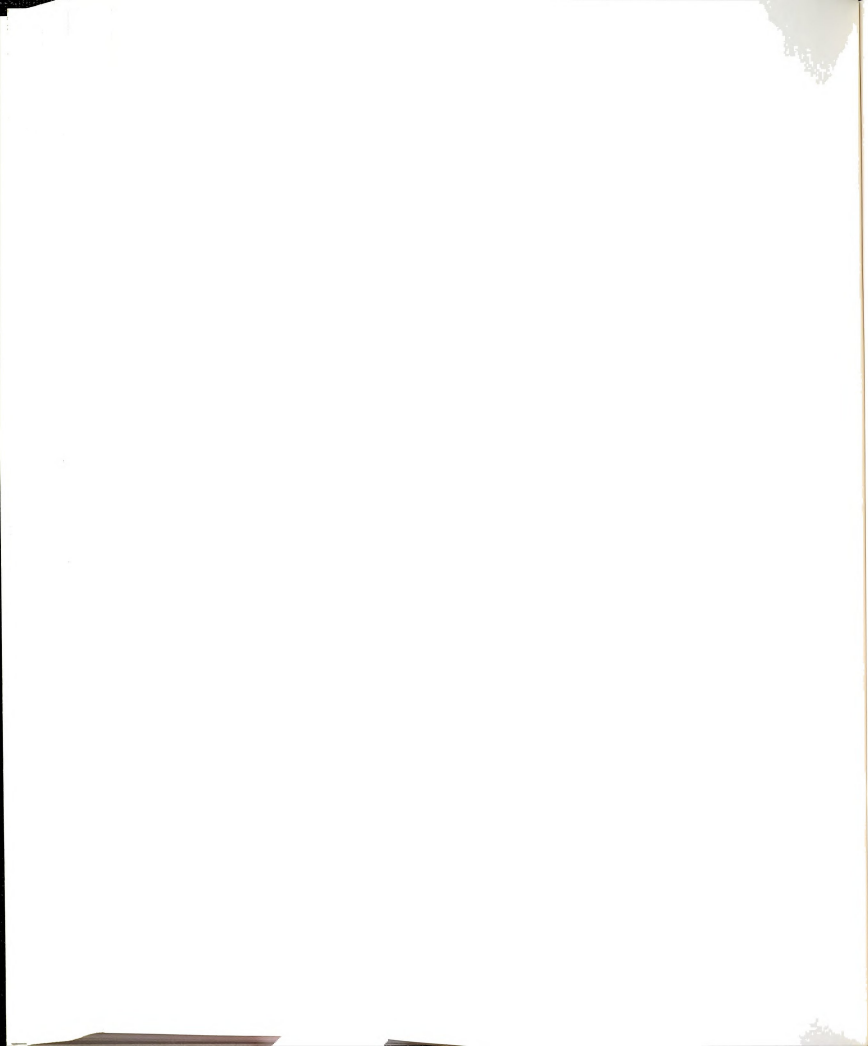
Figure D.6: Histogram of the subjects' responses for problem F. Bayes' estimate denoted by the arrow sign <--.



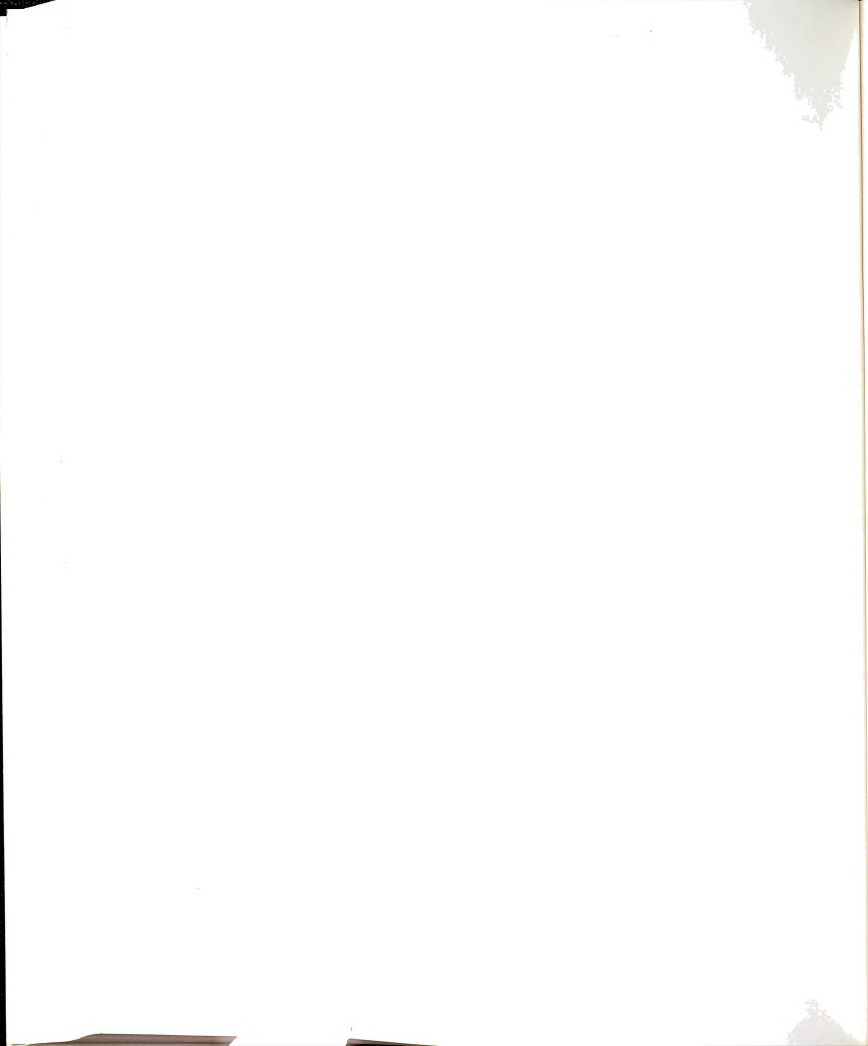
LIST OF REFERENCES

- Ajzen, I. (1977) Intuitive theories of events and the effects of base rate information on prediction. Journal of Personality and Social Psychology, 35, 304-314.
- Bar-Hillel, M. (1980) The base rate fallacy in probability judgment. Acta Psychologica, 44, 211-33.
- Bieri, J.; Atkins, A.L.; Briar, S.; Leaman, R.L.; Miller, H. & Tripodi, T. (1966) Clinical and Social Judgment: The discrimination of Behavioral Information. N.Y. : John Wiley & Sons, Inc.
- Bishop, Y.M.M.; Fienberg, S.E. & Holland, P.W. (1975) Discrete Multivariate Analysis. Cambridge: The MIT Press.
- Borgida E. & Brekke, N. (1981) The base rate fallacy in attribution and prediction. In J. H. Harvey, W. J. Ickes & R. F. Kidd (Eds.) New Directions in Attribution Research (Vol 3). Hillsdale. N.J.: Erlbaum.
- Bruner, J. (1986) Actual Minds, Possible Worlds. Cambridge: Harvard University Press.
- Conover, W.J. (1980) Practical Nonparametric Statistics

- Cohen, L.J. (1981) Can human irrationality be experimentally demonstrated? The Behavioral and Brain Sciences, 4, 317-70.
- Ericsson, K.A. and Simon, H. (1980) Verbal reports as data. Psychological Review, 87, 3, 215-250.
- Evans, J.ST.B.T. (1989) Bias in Human Reasoning. UK: Lawrence Erlbaum Associates Ltd.
- Fischhoff, B., & Bar-Hillel, M. (1984) Diagnosticity and the base rate effect. Memory and Cognition, 12, 402-10.
- Gibbs, G., Morgan, A. & Talor, L. (1980) A review of the research of Ference Marton and the Goteborg Group Institute of Educational Technology, The Open University, Study Methods Group, report no. 2.
- Ginosar, Z. & Trope, Y. (1980) The effects of base rates and individuating information on judgments about another person. Journal of Experimental Social Psychology, 16, 228-42.
- Ginosar, Z. & Trope, Y. (1987) Problem solving in judgment under uncertainty. Journal of Personality and Social Psychology, 52, 464-74.
- Hammerton, M. (1973) A case of radical probability estimation. Journal of Experimental psychology, 101, 252-54.



- Hinsz, V.B., Tindale, R.S. Nagao, D.H., Davis, J.H., & Robertson, B.A. (1988) The influence of the accuracy of individuating information on the use of base rate information in probability judgment. Journal of Experimental Social Psychology, 24, 127-45.
- Holt, R.R. (1958) Clinical and Statistical prediction: A reformulation and some new data. Journal of Abnormal and Social Psychology, 56, 1-12.
- Holt, R.R. (1986) Clinical and Statistical prediction: A retrospective and would-be integrative perspective. Journal of Personality Assessment, 50, 376-86.
- Kahneman, D., Slovic, P. & Tversky, A. (1982) Judgment under Uncertainty: Heuristics and biases. N.Y.: Cambridge University Press.
- Kahneman, D. & Tversky, A. (1973) On the psychology of prediction. Psychological Review, 80, 237-51.
- Kleinmuntz, B. (1990) Why we still use our heads instead of formulas: Toward an integrative approach. Psychological Bulletin, 107, 296-310.
- Lyon, D., & Slovic, P. (1976) Dominance of accuracy information and neglect of base rates in probability estimation. Acta Psychologica, 40, 287-98.
- Marton, F. (1981) Phenomenography- describing conceptions of the world around us. Instructional Science, 10, 177-200.



- Marton, F. & Saljo, R. (1984) Approaches to learning. In F. Marton, D. Hounsell and N. Entwistle (Eds) The Experience of Learning. Edingurgh: Scottish Academic Press.
- Meehl, P.E. (1954) Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. Minneapolis: University of Minnesota Press.
- Meehl, P.E. (1986) Causes and effects of my disturbing little book. Journal of Personality Assessment, 50, 370-75.
- Newell, A. & Simon, H.A. (1972) Human Problem Solving. Englewood Cliffs, N.J.: Prentice Hall.
- Nisbett, R.E. & Ross, I. (1980) Human Inference: Strategies and short-comings and social judgment. Englewood Cliffs: Prentice Hall.
- Nisbett, R.E., Zukier, H., & Lemley, R.E. (1981) The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. Cognitive Psychology, 13, 248-77.
- Strauss, A. (1987) Qualitative Analysis for Social Scientists. Cambridge: Cambridge University Press.
- Tversky, A. & Kahneman, D. (1980) Causal schema in judgment under uncertainty. In M. Fishbein (ed.) Progress in Social Psychology, 49-72, Hillsdale, N.J.: Erlbaum.

White, P. (1984) A model of the layperson as pragmatist.

Personality and Social Psychology Bulletin, 10, 333-48.





MICHIGAN STATE UNIV. LIBRARIES



31293007906914