TO CONSULT OR NOT: A SIGNAL DETECTION STUDY OF ADVICE-TAKING

By

SHENGHUA LUAN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2004

This dissertation is dedicated to my wife Tian Liu, the love of my life.

ACKNOWLEDGMENTS

I would like to thank the members of my doctoral committee for their support and guidance during the qualifying process, and through the development and completion of this dissertation. These include Dr. Robert Sorkin, committee chair; Dr. Ira Fischler and Dr. James Shepperd of the Department of Psychology; and Dr. Joseph Alba of the Department of Marketing. They have all contributed to making this accomplishment an enjoyable and enlightening process. I would also like to thank my wife Tian Liu, a truly intelligent graduate student in Statistics at the University of Florida, for her love, encouragement, and support. I would also like to thank my mother and father in Beijing, China for their years of patience, guidance, and complete support. I am indebted to all of them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

TO CONSULT OR NOT: A SIGNAL DETECTION STUDY OF ADVICE-TAKING

By

Shenghua Luan

August 2004

Chair: Robert D. Sorkin
Major Department: Psychology

Seeking advice from other people or sources is a common practice in making real-life decisions. In this study, a normative model based on the signal detection theory (SDT) was developed to describe how people should acquire and use advice in a detection task. The model shows how advice-taking should depend on the decision-maker's own estimate as well as the cost, expertise, and decision bias of the potential advisors. Based on this model, three experiments were conducted to explore how human participants took and used advice in different decision environments, and whether their observed behaviors were consistent with the model's prescriptions.

In Experiment I, participants were asked to decide whether to consult an advisor' opinion on each experimental trial. Both the cost and the displayed format of the advice were manipulated. It was found that participants consulted more frequently when their own estimates were low, and less frequently when they were high; and they could adjust their consulting frequencies in proper directions in different task conditions. In

Experiment II, two advisors having different expertise and information costs but equal information values were available to consult. Participants were found to consult the "low-expertise-low-cost" advisor more often than the "high-expertise-high-cost" one. This result indicated that participants weighed "cost" more heavily than "accuracy" in the accuracy-cost tradeoff involved in advice-taking. In Experiment III, two advisors with equal expertise but different decision biases were presented. Our model prescribed that a consulting strategy based on a "confirmation-bias-like" rationale would be the best strategy for the participants to use. And this strategy was found to be the one that most participants actually adopted to take biased advice.

Based on the results from those three experiments, it can be concluded that people have certain behavioral tendencies and are task-sensitive when taking advice from external sources. And our SDT model can be useful for identifying whether and how those observed behaviors are deviated from the optimal. With this model as the theoretical guidance, more studies focusing on issues in advice-taking are expected to be conducted in the future.

CHAPTER 1
INTRODUCTION

Seeking advice from other people or sources is a common practice in making real-life decisions; whether as simple as finding directions in an unfamiliar place, or as complex as those involving life-related medical or legal issues. A major motivation of this advice-seeking activity is "people's need to improve their decision accuracy and their expectation that advice will help" (Yaniv, 2004). To achieve this goal, however, at least two important processes need to be considered: determining when and from which sources the needed information should be gathered, and finding how this information should be integrated with the existing information a decision-maker has already had. In this chapter, I will first review literatures related to those two issues, and then present our own integrated model that defines the optimal advice-seeking-utilizing behavior.

Studies in Information Purchasing

Information usually is obtained with a certain cost. It can be the "cost of thinking" (e.g., Russo & Dosher, 1983); the cost of time and its resulting utility loss (e.g., Stigler, 1961); or the real monetary cost (e.g., Connolly & Thorn, 1987). No matter in what form it takes, it needs some sort of investment from the decision maker. Because more information generally leads to more accurate decisions (e.g., Ashton & Ashton, 1985; Sorkin, Hayes, & West, 2001; Yaniv, 1997; but see Wilson & Schooler, 1991), a tradeoff between the expected accuracy improvement and the cost of information acquisition must be considered before (and sometimes during) the advice-seeking process. Studies in information purchasing focused particularly on this issue.

Two major research paradigms had developed studying the topics in information purchasing: the Bayesian paradigm and the Regression paradigm. Although those two paradigms differ in their models defining the optimal tradeoffs and the experimental procedures used in their laboratory studies, they share some common characteristics. First, instead of making external information a given feature in the decision environment, they require a decision maker to obtain it with some measurable (mostly monetary) cost. In addition, a decision payoff function that basically rewards correct responses and penalizes incorrect ones is also embedded in the task. Moreover, the correlation between the amount of information and decision accuracy is generally positive. So, in order to maximize her total gain (decision payoff minus information cost), a decision maker must trade off the improved decision quality, therefore increased decision payoff, with the increased cost. To study how people behave in such tasks was the main research purpose of both paradigms. By chronological order, the Bayesian paradigm will be reviewed first.

As its name indicates, research in the Bayesian paradigm is built largely on the foundation of the Bayes' Theorem. Initiated by the work of some of the earliest decision researchers (Wald, 1947; Raiffa, 1968; Edwards, 1965), it generated a sizable research in the 1960s and 1970s. In a typical experiment under this paradigm, a participant is asked to test one of two (or more) hypotheses by paying to sample the evidence pool; and then judge whether that hypothesis is true. For example, a participant would be first shown two bags and allowed to ascertain the contents of each, say 30 white and 70 black chips in Bag A; 30 black and 70 white chips in Bag B. The experimenter then shuffles the two bags and randomly selects one. The participant's task is to determine which one, by sampling at a fixed cost per chip, the selected bag is: A or B. As a result, she will receive a payoff according to the correctness of her decision. Therefore, testing categorical hypotheses by using discrete data is the main theme of the Bayesian paradigm.

The first question this paradigm tried to resolve concerns the superiority of two stopping rules. In the first, called “fixed stopping”, a decision-maker must specify ahead of time how much data (in the above example how many chips) she wishes to buy. In the second, “optional stopping”, she can buy one chip at a time and may decide at any point when to stop buying information and proceed to make a decision or a posterior probability estimate. With other things equal, it was found that optional stopping is a better strategy than fixed stopping, because it allows the decision-maker to exploit her good luck of gathering some especially diagnostic information in a early stage, hence reducing the purchasing cost (Edwards, 1965).

When a specific stopping rule is set and the related task variables, like the decision payoff function, purchasing cost, and distribution of the samples, are known, an optimal tradeoff point of how much data should be purchased can be calculated by statistical models developed from the Bayes’ Theorem (e.g., Edwards, 1965; Wendt, 1969). Like most models of this nature, “optimality” here is defined as the maximization of the expected monetary value. Those models made the comparisons between the optimal and observed information purchasing behaviors possible; and made the results from laboratory studies meaningful.

Major findings from this paradigm (see reviews like Connolly, 1988; Einhorn & Hogarth, 1982; Peterson & Beach, 1967) include the following. First, people generally respond in the normatively appropriate direction to variations in the task characteristics, like information diagnosticity and cost, but typically less than is normatively justified.

Second, people are inappropriately sensitive to some normatively irrelevant variables. For example, they bought more information when more was simply made available. Third, although few studies explored participants' abilities to improve in information purchasing tasks, slow or no learning was observed in those studies that had. And finally, both over- and under-purchase (compared to the optimal) were found. But it is not clear why both patterns occurred.

The Regression paradigm, represented by the work of Terry Connolly and his colleagues (Connolly, 1977b; Connolly & Gilani, 1982; Connolly & Thorn, 1987; Connolly & Wholey, 1988), followed up the "optimal-model-to-laboratory-experiment" research methodology, but adopted a very different approach to studying information purchasing. Instead of dealing with decision problems with discrete hypotheses and data, its tasks involved judgment or estimation of a continuous variable on the basis of several continuous cues. Correspondingly, the normative treatments in this paradigm are drawn from regression models rather than the Bayes' Theorem. Therefore, it shares lots of similarity with the tradition of the Brunswikian School's lens model and the research from the multiple cue probability learning (MCPL) (Brehmer & Joyce, 1988).

In a typical experiment under this paradigm, a participant would be asked to make a prediction or estimate, say the results of a series of football games, based on the value or values of one or more cues, such as assessments from a bunch of "football experts". To obtain each cue's value, however, she would need to pay a certain cost. And the errors of her final assessments would be penalized according to a squared-error loss function. The causal relation between the estimated variable (or the "criterion" variable in the lens model's term) and those cue variables may be different among tasks. But in each task condition, by specifying the payoff function, cues' validities, and the purchasing cost(s), the optimal purchasing behavior of which cues and how many of them should be bought can be worked out by some Regression models.

The results from laboratory studies of the Regression paradigm are similar to those found in the Bayesian paradigm: Although people are somewhat responsive to the variations of the task variables, their adjustments across different task environments are usually not sufficient. They respond strongly to normatively irrelevant factors; and have great difficulty assessing the validity of the cues they are offered. They sometimes under-purchase cue information and sometimes over-purchase. However, unlike in the Bayesian paradigm, a learning mechanism based on the trial-and-error strategy was found to be able to explain people's purchasing pattern fairly well (Connolly, 1988).

Both the Bayesian and regression analyses have earned enormous popularity in the field of judgment and decision-making as the normative standards to treat information. Studies in information purchasing give an example how they can be used to deal with the accuracy-cost tradeoff problems concerning people's need to have more information and its associated cost. This research addresses an important yet often ignored topic in human decision making: pre-decisional information acquisition; and contributes fruitful insights to the processes involved. However, the general finding is pessimistic: people are far from being optimal in their acts to purchase information and usually perform poorly at balancing the value of the information and its cost. This is perhaps the most important reason why this line of research has almost vanished after the 1990s. While information purchasing was left behind, research in advice-taking has tried to gain its influence from the late 1990s; and I will give a brief review of findings in that area next.

Studies in Advice-Taking

Integrating acquired external information with existing information seems like a natural step following information purchasing. Although it indeed is the necessary part of all normative models in information purchasing, information integration seldom makes the center of the issues. But it is the focus in the advice-taking studies. With no interest in how people manage to acquire external information, “advice” or source information is usually a “given” in the task environments of those studies; and trying to understand how people actually use the advice is the drive of this research.

Information integration is hardly a new topic in the decision literature. Different integration models can be found in domains like attitude change (Anderson, 1968; Zimbardo & Leippe, 1991), combination of expert opinions (Budescu & Rantilla, 2000; Clemen & Winkler, 1999), social judgment (Brehmer & Joyce, 1988), and group interactions (Kerr, MacCoun, & Kramer, 1996; Sorkin et al., 2001). The difference between the type of information integration in advice-taking and the others’ is subtle. Before taking any advice, the decision maker has usually formed her own opinions or judgment. Information from those advisors is made available to use but how to use it is totally up to the decision maker. Researchers are interested in how one’s initial opinions are shifted by the opinions from the advisors when her final opinions are formed; and the cognitive and social factors that may influence this process of advice using. The advice recipient is the focused entity during the whole process, and her own initial opinions are an important task variable, just as the advisors’. These characteristics are shared by all advice-taking studies, despite some variations designed to fit each study’s purposes.

In the judge-advisor paradigm proposed by Sniezek and Buckley (1995), the authors studied a particular advice-taking system in which two advisors formulated their decision estimates and communicated the information to a person who was responsible for making a final decision, the judge. They found in a series of experiments that whether a judge had formed her independent opinions before seeing any advice affected her final decision accuracy and confidence, as well as her using of the advice. Also, advisors' stated confidence and the decision conflict between their opinions were found to affect how their advice would be used by the judge, with more influence from the more confident advisor and less influence in overall when the decision conflict was high. Based on these findings, the authors concluded that advice-taking is a dynamic system in which both the judge's initial opinions and the characteristics of the advisors' opinions are critical in shaping the judge's advice-taking behaviors.

To further understand the judge's role in this dynamic judge-advisor system, Harvey and Fischer (1997) conducted their own study trying to answer a fundamental question: why do people take advice? They offered three reasons. First, the very existence of other people offering advice puts people under social pressure to comply with it. People are willing to accept help from advisors and revise their opinions, sometimes even when those advisors' information has no diagnostic value. Second, people may hold the belief that by taking advice, they may be able to share responsibility for important judgments and decisions with the advisors, so that they could avoid more severe punishment of the possible negative consequences. Third, people accept advice simply because it improves their decision or judgment accuracy, as many other studies also found. Alongside exploring why people take advice, the authors also found that people are sensitive to the expertise level of the advisors and her own. The perceived expertise is an important factor in determining how much influence the advisors' opinions can have on the judge; and their effects are mediated by the perceived importance of the undergoing task.

To integrate advice with one's own estimate, a weighing function with decisional weights assigned to the opinions from each source seems necessary. Understanding the characteristics of this weighing process has been the main goal of Ilan Yaniv's advice-taking research (Yaniv, 2004; Yaniv & Kleinberger, 2000). In their studies, they found that people are often biased to place a higher weight on their own opinions than on the advisors', a phenomenon the authors called the "self/other effect". Therefore, advice, even of high diagnosticity, is often discounted by its recipients. They found also that people tended to discount inconsistent advice. That is, the bigger the perceived disagreement between an advisor's opinion and the decision-maker's own, the smaller the weight of that piece of advice receives, a "distance effect" as called by the authors. These two effects lead to the sub-optimal use of advice, although it still helps improve the final accuracy significantly.

Advice-taking is a very practically useful yet theoretically less developed area. Besides the empirical findings listed above, there still does not exist a formal model in the area. Also, considering studies in information purchasing, the assumption of "free" advice may not be as practical as it is in real-life situations. To fully examine the whole process of advice-taking, adding a pre-integration stage of information acquisition seems both theoretically and empirically necessary. To remedy those problems, we developed our own model of advice-taking based on the signal detection theory. It integrates the important characteristics from both the information purchasing and existing advice-taking studies, and has its own definition of normative behaviors and task procedures. In the next session, I will give a brief introduction of this model and describe how it can be applied to study both information purchasing and advice-taking behaviors as a whole.

A SDT Model of Advice-Taking

The signal detection theory has been best known for its applications in perception studies in psychology (Green & Swets, 1966). However, developed from the field of Electrical Engineering, its original purpose was to understand how the input information can be optimally processed by one or a bunch of signal detectors. Sensing the similarity between hardware signal detectors and individuals in a group setting, Sorkin and Dai (1994) first introduced SDT as an approach to group decision studies. After that and a series of other studies (summarized in Sorkin, Luan, & Itzkowitz, in press), the SDT method has been proved not only useful to examine human decision behaviors in a group setting, but also a promising method to study information integration at the individual level (e.g., Luan, Sorkin, & Itzkowitz, 2004). Because of the success of those early work, we believe that we can model advice-taking by using SDT also.

A model can be useless without specifying the range of tasks it can be applied to. Therefore, I will first describe what a SDT task in advice-taking may look like before introducing the formal mathematical deductions of the model. As shown in Figure 1-1, a decision maker (DM) will first observe the occurrence of a certain decision event. That event can be anything that will lead a DM's initial decision estimate to be correct with a certain probability. In present study we choose to display to the DM a visual stimulus with controlled statistical properties. It can be either a "Signal" or a "Noise" event in each experimental trial. (I will explain the definitions of those two events later.) By doing this, we hope to better control the decision accuracy of the DM's own estimate, and maintain a small variance among participants.

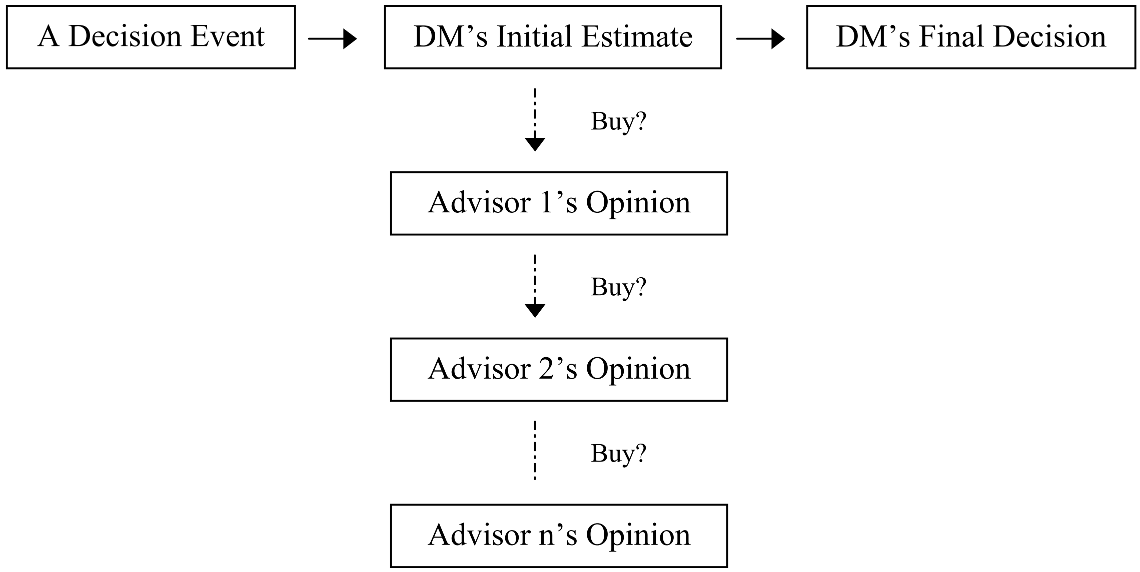

Figure 1-1. Structure of our SDT advice-taking task.

After the visual display, the DM will be asked to report her initial decision estimate regarding the occurrence of the event in the trial. This estimate can be either a binary decision or a continuous likelihood rating. (But we assume that a DM can always have a rating estimate even when only a binary decision is asked to report.) Then, she will be given the chance to consult some "experts", or advisors, about the decision she is about to make. The advisors' information can also be either a binary decision or a continuous likelihood rating, depending on the manipulation of a particular task. In present study, those advisors are not real people, and neither is their information recorded from real people's responses. Their information is computer-simulated with certain controlled statistical properties. And as in those information purchasing studies, there is a certain monetary cost associated with each piece of advice. The DM can choose to buy a piece of information from any advisor and can buy any pieces that are available.

Finally, when the DM finishes her purchase, she will be asked to integrate the obtained information with her own estimate to make a final decision. This final decision will be a binary decision indicating whether the displayed stimulus is a "Signal" or a "Noise". There will be a payoff related to the correctness of the decision that rewards correct decisions and penalizes incorrect ones; and it can be symmetric for rewards and penalties or asymmetric. A DM's only purpose in such a task is to earn as much money (or equivalent game points) as possible by carefully spending her money in purchasing advice and finding the best way to combine multiple pieces of information.

From the description of the task, we can see that there are two main processes: information acquisition and integration, involved in the task. Based on SDT, our model can define the optimal behaviors in each process. To explain how, we will begin with a very simple task: When there is only one advisor's information available to consult. After working this case out, it should not be hard to expand the model to more complex tasks.

When there is only one advisor available, the most important question for a DM in the acquisition stage is: "Should I buy the advisor's information?" An intuitive answer to this question is that it should depend on the certainty, confidence, or likelihood estimate of the DM's own observation. If it is high toward one particular decision, she should not buy the information; and if it is low, she should buy it. In order to do this, a DM should set up a criterion in her mind to determine whether her initial estimate is high or low, so that she can choose the corresponding action. Given this is the case, in what value should this criterion be? Is there an "ideal buying criterion" by adopting which the DM's expected value can be maximized? The answer is yes; and as our model indicates, the selection of this ideal criterion should depend on several important task variables. Before the illustration of our model, however, we think that it is better for us to briefly introduce some basic concepts in SDT, so that readers can have a better understanding how the model works.

In SDT, an observation of the input stimulus by the DM can be represented by an x value to indicate its magnitude. It is assumed that for each of the two decision events: "Noise" and "Signal", there is a mapping probability distribution of x: f(x/Noise) and f(x/Signal), in the DM's observations. And those two distributions are usually assumed to be normally (Gaussian) distributed with equal variances (σ^2) but different means (μ), as illustrated in Figure 1-2. A DM's ability to discriminate between these two events is summarized by the difference between the means of the distributions divided by their standard deviation, $(\mu_1 - \mu_0)/ \sigma$. This normalized separation between the means is termed as the detection index, d' (Tanner & Birdsall, 1958). Therefore, if two different DMs have the same Noise distribution, the difference in the means of their Signal distributions can tell their difference in detectability or decision accuracy.

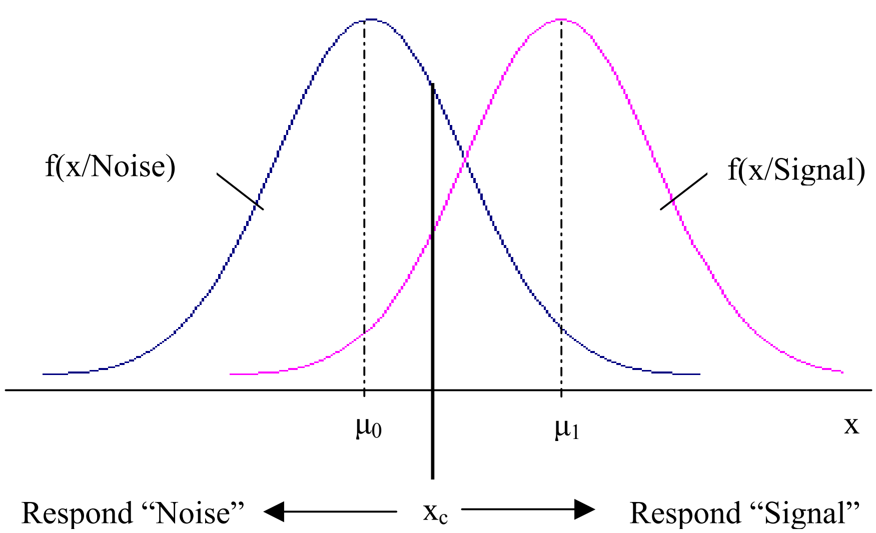

Figure 1-2. Decision statistics in a SDT task.

How should a DM make a decision without any advice? It has been shown that if a DM can convert her observation x to a likelihood ratio, $\lambda(x)$, where

$$\lambda(x) = \frac{f(x/Signal)}{f(x/Noise)} \qquad 1\text{-}1$$

and makes her decision according to this statistic, her decision to be made can be optimal by a number of different criteria, including maximizing expected value (Macmillan & Creelman, 1991). Suppose that a DM can make such conversions; then there must be a criterion value above which the DM will respond a "Signal" and under which she will respond a "Noise" in $\lambda(x)$. And this value is termed as the response criterion, β, in SDT (Green & Swets, 1966).

Two things need to be mentioned about β: First, for a particular criterion β, there should be a matching x, which is called x_c, in observation; and $\beta = \lambda(x_c)$. Second, if the priori odds P(Signal)/P(Noise)=1 and the rewards for correct decisions are the same as the penalties for incorrect ones, a DM should adopt a response criterion at $\beta_{DM} = 1$ (Green & Swets, 1966). If her β is found to be bigger than 1 ($\beta_{DM} > 1$), it indicates that the DM has a tendency to make more Signal decisions; and if $\beta_{DM} < 1$, it indicates that the DM tend to make more Noise decisions.

When information from multiple sources is available, how should a DM integrate it with her own? Suppose there are totally n+1 sources, including the DM; and all sources' information is expressed in the form of a likelihood ratio: $\lambda(x)$. Therefore, we have $\lambda(x_0)$, $\lambda(x_1)$, $\lambda(x_2)$, …, $\lambda(x_n)$ as the information input from the DM, source 1, source 2, ..., and source n, respectively. The DM's task is to update her own likelihood ratio estimate $\lambda(x_0)$ with other likelihood ratios to form an integrated statistic, $\lambda(x_{0\ldots}x_n)$; and makes a decision based on its value. When all sources are independent, the optimal way to calculate this statistic is simple:

$$\lambda(x_0...x_n) = \frac{f(x_0...x_n / Signal)}{f(x_0...x_n / Noise)} = \frac{f(x_0 / Signal)...f(x_n / Signal)}{f(x_0 / Noise)...f(x_n / Noise)} = \prod_{i=0}^{n} \lambda(x_i) \qquad \text{1-2}$$

Therefore, if there is only one advisor (A) available and A's observation is made independently from the DM's, the integrated likelihood ratio estimate $\lambda(x_0x_1)$, or λ(DM-A), is simply the multiplication of the two likelihood ratios.

When there is a correlation among the sources, the optimal integration calculation will become more complicated. And its solution can be found in somewhere else (e.g., Durlach, Braida, & Ito, 1986). If sources' information is not expressed in the form of a likelihood ratio but some external representation of it, say a likelihood rating L(x) in a 0-100 scale (as used in present study), a DM should form another statistic, say $L(x_{0...}x_n)$, and make a decision according to its value. It can be deduced that the optimal calculation of $L(x_{0...}x_n)$ is to form a linear combination of the ratings with weights proportional to sources' d's (Sorkin & Dai, 1994):

$$L(x_0...x_n) = d'_0 L(x_0) + ... + d'_n L(x_n) \qquad \text{1-3}$$

After a review of the important statistics in SDT and how multiple pieces of information can be optimally integrated in a SDT task, let us return to the exemplary case where only one advisor's information is available to purchase. As discussed before, a DM should set up a "buying criterion" to determine when to buy the information and when not to; and this criterion should be a cutoff value in her stated likelihood rating, L(x). Because L(x) is just the external representation of λ(x) and should be in some sort of function of λ(x), we can use λ(x) to substitute L(x). Therefore, the buying criterion can be represented by a λ(x) value. However, because λ(x) is a function of x, we can also represent the buying criterion in the term of the observation, x. The relationships between x, λ(x), and L(x) can be more clearly seen in Figure 1-3.

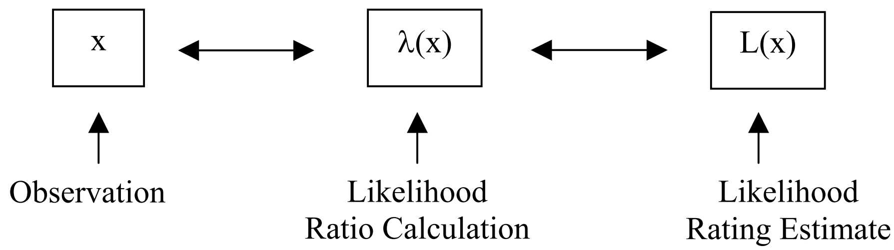

Figure 1-3. Relations between x, λ(x), and L(x).

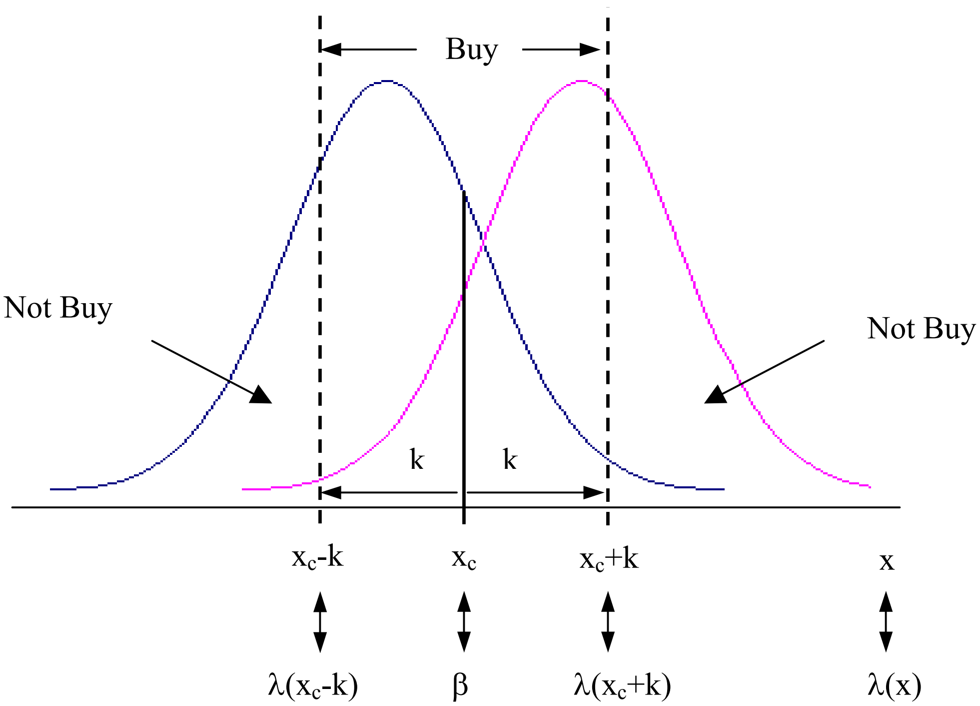

Figure 1-4. Hypothesized information buying process of our SDT model.

The hypothesized buying process proposed by our model is illustrated in Figure 1-4. Suppose a DM has an observation, x, and needs to make a decision whether it is a Signal or Noise. The model says that: 1) If the observation x is bigger than x_c+k ($x > x_c+k$), where k is a constant, a DM will not buy the information from the advisor; and she will make a Signal final decision directly. 2) if x is in the range: $[x_c, x_c+k]$, she will buy the information. Furthermore, she will not make her final decision until after seeing the advice. 3) if x is in the range: $[x_c-k, x_c]$, she will also buy the information and wait after seeing the advice to make a final decision. 4) if $x < x_c-k$, the DM will not buy the information; and she will proceed to make a Noise final decision.

Therefore, whenever the DM has a strong observation, whether it is towards the Signal or Noise decision, she should not buy the advice from the advisor (A). She should buy the information only when her observation is weak, where the Signal and Noise distributions heavily overlap so that it is harder to tell them apart. And $\lambda(x_c+k)$ and $\lambda(x_c-k)$ should be her buying criteria in the Signal and Noise distributions respectively. To maximize her expected value, EV, the DM should select a proper "k" value as the determination of her ideal buying criteria. How to calculate an EV with a certain k is illustrated in the following. And as in many other mathematical deductions, we will begin with the assumptions and parameters that will be used.

Assumptions

- The prior odds ratio is equal to 1: P(Signal)/P(Noise)=1.
- f(x/Noise) for both the DM and A is a standard Normal distribution: N(0, 1).
- f(x/Signal) for the DM is a Normal distribution: $N(\mu_1, 1)$. Therefore,
- $d'_{DM} = (\mu_1-0)/1= \mu_1$.
- f(x/Signal) for A is a Normal distribution: $N(\mu_2, 1)$. Therefore, $d'_A= \mu_2$.
- DM's criterion is: $\beta_{DM}=1$, or $x_{c0}= \mu_1/2$.
- A's criterion is: $\beta_A=1$, or $x_{c1}= \mu_2/2$.

- DM's criterion does not change after integrating A's information.
- Both DM and A's information is expressed in $\lambda(x)$.

Parameters

- k: the shift of x_{c0} in both distributions and is represented by a *z* score.
- EV: the expected value in each trial of the task.
- C: information cost, the cost of buying A's information.
- V: the decision payoff. It is symmetric for the reward and penalty, and is a positive value for reward (+V) and a negative value for penalty (-V).
- P(Buy): the probability that the DM buys the information from A.
- P(Correct): the probability that the decision made is correct, in both conditions when the DM buys the advice and does not buy.
- P(Incorrect): the probability that the decision made is incorrect, in both conditions when the DM buys the advice and does not buy.

Calculations

- **P(Buy):**

$$P(Buy) = P(x_{c0} - k \le x \le x_{c0} + k) = P(\frac{\mu_1}{2} - k \le x \le \frac{\mu_1}{2} + k)$$
$$= P[(\frac{\mu_1}{2} - k \le x \le \frac{\mu_1}{2} + k) / Noise] + P[(\frac{\mu_1}{2} - k \le x \le \frac{\mu_1}{2} + k) / Signal]$$
$$= P(Noise) * [\phi(\frac{\mu_1}{2} + k) - \phi(\frac{\mu_1}{2} - k)] + P(Signal) * [\phi(-\frac{\mu_1}{2} + k) - \phi(-\frac{\mu_1}{2} - k)]$$
$$= \phi(\frac{\mu_1}{2} + k) - \phi(\frac{\mu_1}{2} - k)$$

In this equation, Φ(x) represents the integration from the negative infinity to x in a standard Normal distribution: the area under x in N(0, 1).

- **P(Incorrect):**

There are three additive parts in P(Incorrect), represented by P(a), P(b), and P(c). P(a) is the probability that the decision made by the DM *without* consulting the advisor is incorrect. P(b) is the probability, when the stimulus is Noise, that the DM makes an incorrect decision *after* consulting. P(c) is the probability, when the stimulus is Signal, that the DM makes an incorrect decision *after* consulting. Because the prior odds is equal to 1, P(b) is equal to P(c).

  - **P(a):**

$$P(a) = P(x > \frac{\mu_1}{2} + k) * P(Noise) + P(x < \frac{\mu_1}{2} - k) * P(Signal)$$
$$= 1 - \phi(\frac{\mu_1}{2} + k)$$

  - **P(b) and P(c):**

In the Noise distribution, "b" happens when x is in the area: $\mu_1/2-k < x < \mu_1/2+k$, *and* the updated likelihood ratio $\lambda_{DM}\lambda x_A>1$. Suppose that x is the observation of the DM and y is the observation of A. From the ideal integration $\lambda_{DM}\lambda x_A >1$, we get:

$$\lambda_{DM} * \lambda_A > 1$$
$$\Rightarrow Ln\,\lambda_{DM} + Ln\,\lambda_A > 0$$
$$\Rightarrow x\mu_1 + y\mu_2 - \frac{\mu_1^2 + \mu_2^2}{2} > 0$$
$$\Rightarrow y > (\frac{\frac{\mu_1^2 + \mu_2^2}{2} - x\mu_1}{\mu_2})$$

We can use this deduction in the following calculation:

$$P(b) = P(N) * P(\frac{\mu_1}{2} - k \le x \le \frac{\mu_1}{2} + k; \beta_{DM} * \beta_A > 1 | x \in N)$$

$$= P(Noise) * \int_{\frac{\mu_1}{2}-k}^{\frac{\mu_1}{2}+k} \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}) \int_{\frac{\frac{\mu_1^2+\mu_2^2}{2} - x\mu_1}{\mu_2}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dydx$$

$$= P(Noise) * \int_{\frac{\mu_1}{2}-k}^{\frac{\mu_1}{2}+k} \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}) * [1 - \phi(\frac{\mu_1^2 + \mu_2^2}{2\mu_2} - \frac{x\mu_1}{\mu_2})] dx$$

$$= P(Noise) * \left\{ \phi(\frac{\mu_1}{2} + k) - \phi(\frac{\mu_1}{2} - k) - \int_{\frac{\mu_1}{2}-k}^{\frac{\mu_1}{2}+k} [\frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}) * \phi(\frac{\mu_1^2 + \mu_2^2}{2\mu_2} - \frac{x\mu_1}{\mu_2})] dx \right\}$$

Because the integration part in the right-side looks very complicated, we use D(k) to represent it. And because P(Noise)+P(Signal)=1, and P(b) is equal to P(c), therefore:

$$P(b) + P(c) = \phi(\frac{\mu_1}{2} + k) - \phi(\frac{\mu_1}{2} - k) - D(k)$$

Therefore,

$$P(Incorrect) = P(a) + P(b) + P(c)$$
$$= 1 - \phi(\frac{\mu_1}{2} + k) + [\phi(\frac{\mu_1}{2} + k) - \phi(\frac{\mu_1}{2} - k) - D(k)]$$
$$= 1 - \phi(\frac{\mu_1}{2} - k) - D(k)$$

- **P(Correct):**

Because P(Correct) + P(Incorrect) =1,

$$P(Correct) = 1 - P(Incorrect) = \phi(\frac{\mu_1}{2} - k) + D(k)$$

- **EV:**

After working out those probabilities, we can form a linear combination of EV that includes the expected value of making a correct decision, the expected value of making an incorrect one, and the expected cost of buying A's information. The final product should be a function of k, V, C, and the decision characteristics of the DM and A.

$$EV = P(Correct) * V - P(Incorrect) * V - P(Buy) * C$$
$$= [\phi(\frac{\mu_1}{2} - k) + D(k)] * V - [1 - \phi(\frac{\mu_1}{2} - k) - D(k)] * V - [\phi(\frac{\mu_1}{2} + k) - \phi(\frac{\mu_1}{2} - k)] * C$$
$$= \phi(\frac{\mu_1}{2} - k) * (2V + C) - \phi(\frac{\mu_1}{2} + k) * C + D(k) * 2V - V$$

It can be seen that the EV equation is mathematically quite complex. However, in a given task situation, all parameters but k are constants. Therefore, by substituting those constants in the EV equation, we can plot EV as a function of k, and then locate the k value that can maximize EV. With this ideal "k" known, we can subsequently determine the ideal buying criteria.

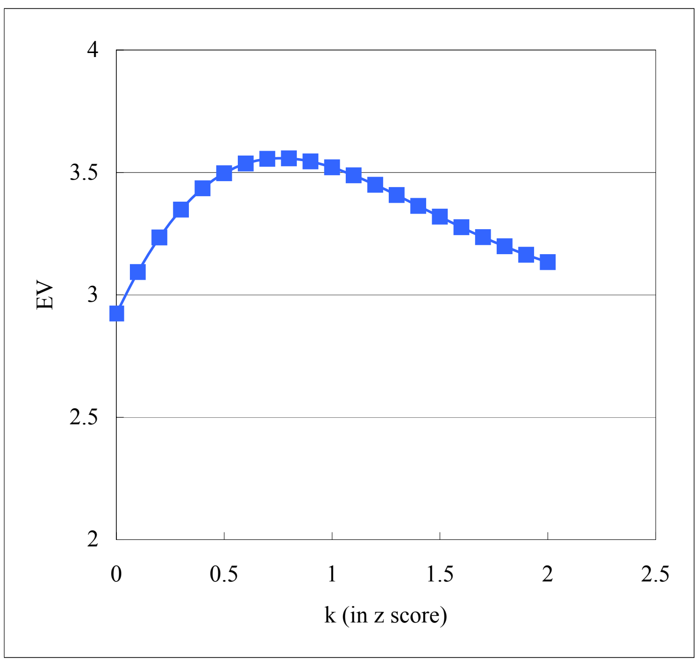

Figure 1-5. EV as a function of k when d'_{DM}=0.75, d'_{A}=1.50, V=10, and C=3.

Figure 1-5 gives an example of such calculations by using those task variables: d'_{DM}=0.75 (μ_1=0.75), d'_{A}=1.50 (μ_2=1.50), V=10 money units, and C=3 money units. It can be seen that EV approaches its maximum when k is around 0.80. Therefore, a k value of 0.80 can be identified as the ideal k. Because k is in the unit of a *z* score, the cumulative probability of the buying area [x_c-k, x_c+k] can be easily calculated. In this case, it is 0.54. It means that if a DM adopts a buying criteria converted from k=0.80, she would buy the advisor's information 54% of the time.

When the number of advisors is more than one and the decision properties of the DM and those advisors as not as simple as assumed in our exemplary case, the mathematical deductions of the EV equation can become even more complex. To deal with this problem, a methodologically simpler but equally effective method: computer simulations, can be used to plot the EV function to locate the ideal k. In present study, most of our normative predictions would be coming from the results of simulations. Because those simulations followed the same principles of our model, their results are just as valid as what might be calculated from the mathematical deductions.

Based on our model, three experiments were conducted to probe the advice-taking behaviors of human participants in a signal detection task. In Experiment I, we manipulated the information cost and the format of an advisor's information to test how people would purchase information under different task environments. In Experiment II, we presented participants with two advisors who had different expertise (d') as well as different information costs but equal information values. The main purpose was to examine which aspect would be weighed more heavily by the participants during the tradeoff: accuracy or cost. Finally, in Experiment III, two advisors with the same detectability but different response criteria (β) would be available for the participants to consult. To probe what kind of consulting strategy or strategies they would adopt under such task was the main goal of that experiment.

CHAPTER 2
EXPERIMENT I

In this experiment, we put a DM in a very simple advice-taking situation: when there was only one advisor she could consult (buy information from). And two variables were manipulated in this experiment: the information cost (Cost) and the format of the advisor's information (Mode). Based on the principles of our model, we first run computer simulations to determine what the normative buying behavior should be in each task condition; and then conducted human experiments to examine how our participants actually behaved in the task. We were interested in whether people could take advice as our normative prescribed; and if not, how and in what direction they would be deviated from the normative. Moreover, we were also interested in how people would adjust their advice-taking behavior in different task situations.

As discussed in Chapter 1, buying advice can be treated as a cost-accuracy tradeoff process. And manipulating related variables from both sides of the tradeoff is a common method to study this process (e.g., Weber, Baron, & Loomes, 2001). This is the main reason why we chose to manipulate Cost and Mode in this experiment. In this study, Cost was simply the monetary cost a DM had to pay to obtain the advisor's opinions. It can be relatively low or high compared to the final decision payoff. Mode represents the way the advisor's information is expressed to the DM. In a SDT task, it can be either a binary decision or a continuous rating. Although it is obvious that Cost represents the "cost" side of the tradeoff, the relation between Mode and accuracy may not be that direct. To clarify this issue, let us begin with how information was simulated in our SDT task.

Simulations

It should be made clear that there were two kinds of simulations in the experiment. One was designed for predicting the normative behavior based on our model, "model-simulations"; and the other was for simulating the information from a computer-created advisor, "experiment-simulations". The advisor's information was essentially simulated in the same way in both kinds of simulations. However, because human participants were recruited to do the experiment, a DM's simulated information was needed only in the model-simulations. And we assume that a DM would always have a continuous rating estimate after making an observation, no matter whether she is asked to express it or not. Therefore, in model-simulations, only the continuous information of a DM was simulated; and it was simulated by the same method as how an advisor's continuous information was done, except for using different decision parameters.

The continuous information of an advisor was simulated by the following method. First of all, a random number, x, was generated from a normal distribution, $N(\mu, \sigma)$, to represent a particular observation of the advisor. When the decision event was a Signal (S), $\mu_1=5$; and when it was a Noise (N), $\mu_0=4$. The two distributions shared a common standard deviation, σ; and the value of σ was determined by the wanted d' of the advisor: $\sigma =(\mu_1-\mu_0)/d' =1/d'$. For example, if $d'_A=1.50$, $\sigma=0.67$; and if $d'_A=0.75$, $\sigma=1.33$. In this way, we could accurately control the d' of an advisor. And in all simulations of this study, the prior probabilities of the S and N events were equal: P(S)=P(N)=0.50.

After those distributional parameters and the value of x were known, the mass probabilities of x being drawn from the two distributions: f(x/S) and f(x/N), could then be calculated. With these two probabilities at hand, the likelihood ratio of x being a S now could be obtained: $\lambda(x)= f(x/S)/f(x/N)$. This likelihood ratio $\lambda(x)$ represented an advisor's continuous information in the model-simulations. However, because $\lambda(x)$ can theoretically range from negative infinity to positive infinity, it is an unrealistic representation of a real advisor's continuous information. To solve this problem, we use a mathematically equivalent form of $\lambda(x)$: the posteriori probability of x being drawn from S: P(S/x), as a simulated advisor's information in the real experiment. When the prior odds ratio is equal to 1(P(S)/P(N) =1), the following relation of $\lambda(x)$ and P(S/x) is true (Green & Swets, 1966):

$$P(S/x) = \frac{\lambda(x)}{\lambda(x)+1} \qquad \text{2-1}$$

Therefore, after converting $\lambda(x)$ by Equation 2-1, the continuous information of a simulated advisor was a variable in a 0-100 scale; and none of the information from the original observation, x, was lost by those transformations in between.

The binary decisions of the simulated advisor were generated on the basis of the continuous information $\lambda(x)$, together with a certain response criterion: β_A. Because β_A would be a constant value in this experiment (β_A=1), we simply regarded an advisor's decision as a S if $\lambda(x) > \beta_A$; and a N if $\lambda(x) < \beta_A$. We can see from this generation process that some information from the observations was lost. For example, a higher likelihood ratio based on a stronger observation, say $\lambda(x)$=5, would be no different from a lower likelihood ratio based on a weaker observation, say $\lambda(x)$=2, in terms of the advisor's decisions. How should a DM treat those binary decisions with degraded information value and integrate them?

Because there are two decision events, S and N, four decision outcomes are possible: "Hit", "Miss", "Correct Rejection", and "False Alarm". They can be represented by four conditional probabilities (rates of their occurrence) as seen in Table 2-1. When the d' and β of an information source are known, the values of those probabilities are fixed and can be calculated (Sorkin et al., in press).

Table 2-1. Definitions of the four possible decision outcomes.

		Decision Event	
		Signal (S)	Noise (N)
Decision Made	Signal (s)	Hit: P(s/S)	False Alarm: P(s/N)
	Noise (n)	Miss: P(n/S)	Correct Rejection: P(n/N)

Based on those probabilities, two likelihood ratios can also be calculated by Equation 2-2 and 2-3. One of them represents the likelihood ratio that can be inferred from a "s" response by the advisor: λ(x/s); and the other represents the likelihood ratio that can be inferred from a "n" response by the advisor: λ(x/n).

$$\lambda(x/s) = \frac{P(s/S)}{P(s/N)} = \frac{P(\text{Hit})}{P(\text{False Alarm})} \qquad \text{2-2}$$

$$\lambda(x/n) = \frac{P(n/S)}{P(n/N)} = \frac{P(\text{Miss})}{P(\text{Correct Rejection})} \qquad \text{2-3}$$

Therefore, a DM can update her likelihood ratio just like in the continuous information case by using one of those two λ(x) values; and which one should be used will depend on the specific response from the advisor. Moreover, those two likelihood ratios can be transformed to two posteriori probabilities through Equation 2-1, in case posteriori probabilities are what the DM needs to integrate advisor's information. The optimal information integration methods have been summarized in Equation1-2 and 1-3 in Chapter 1.

From above descriptions, it can be seen that binary decisions generally contain less information than continuous ratings. Therefore, the final decisions made by a DM with the help from binary information would generally not be as accurate as they would be from the continuous information. And if a DM decides to buy binary decisions from an advisor, the information may not be as useful to her as continuous ratings, when the buying costs are equal. We hope that could explain the relation between the Mode of the advice and the accuracy side of the tradeoff.

Simulation Results and Experiment Design

In the model-simulations of this experiment, the following task variables were controlled to be constant: the accuracy of the DM's initial estimate d'_{DM} and her decision criterion β_{DM}, the accuracy of the advisor's information d'_A and its decision criterion β_A, the information correlation between the DM and the advisor $\rho_{DM\text{-}A}$, and the final decision payoffs: D-payoffs (in token-points that could be converted to money). The values of those controlled variables are listed in Table 2-2.

Table 2-2. Values of the controlled variables in Experiment I.

d'_{DM}	β_{DM}	d'_A	β_A	$\rho_{DM\text{-}A}$	D-payoffs (points)	
0.75	1	1.50	1	0	Hit: 10	False Alarm: -10
					Miss: -10	Correct Rejection: -10

Because there were only two levels of Mode available: Binary and Continuous, both of them were used in the simulations and the experiment. However, because there were potentially infinite values of Cost, we must choose the ones that were more meaningful for our experiment. After trying Cost=1, 2, 3, 4, and 5 in both Mode conditions, the results of the "ideal ks" and their resulting EVs (the MaxEVs) in each condition are listed in Table 2-3. In each condition, we ran 10 Monte Carlo simulations with 10,000 trials in each (5,000 each for S and N events) and averaged the individual results from each run as the final result. And in those simulations and any other model-simulations throughout the rest of this study, we assumed that a DM could integrate information from the advisor optimally (see Equation 1-2 and 1-3).

Table 2-3. Ideal ks and MaxEVs in each Cost by Mode condition in the model-simulations of Experiment I.

	Ideal k		MaxEV	
Cost	Continuous	Binary	Continuous	Binary
1	1.8	1.2	5.01	4.75
2	1.2	1.0	4.19	4.06
3	0.8	0.7	3.56	3.50
4	0.4	0.4	3.14	3.10
5	0.1	0.1	2.94	2.93

It can be seen that: First, when Cost increases, the value of the ideal k decreases. Referring to Figure 1-4, it means that the size of the buying area should be smaller when advice becomes more expensive to get. Thus, a DM should buy from the advisor less frequently when Cost is high. This main effect of Cost holds in both the binary and continuous Mode conditions. Second, there is an interaction effect between Cost and Mode. When Cost is low, the ideal ks in the binary Mode are relatively lower than the ones in the continuous Mode. However, when Cost is high, there is no difference in ideal k between the two information Modes. We can infer from this effect that: A DM should buy advice more frequently when it is a continuous rating than when it is a binary decision, given the Cost for accessing each kind of information is equally low. However, when Cost is high, a DM should buy the advice at approximately same frequency regardless of the Mode of the information. Finally, the similar two effects in ideal k can be also seen in MaxEV: When Cost increases, MaxEV decreases in both Mode conditions; and there is an interaction effect between Cost and Mode in MaxEV.

Based on those simulation results, we selected Cost=1 and Cost=4 as the two levels of Cost we would use in the experiment, mainly because of the experimental effects they might have on a DM's selection of her buying criterion (ideal k). Moreover, because ideal k is a *z* score in our model and the above simulations, a more direct but equally effective measure of a DM's buying behavior in the experiment would be her frequency of consulting: the percentage of times she consulted the advisor in a block of trials. This "consulting frequency" measure is the same as the cumulative probability (size) of a buying area defined by an ideal k. Therefore, when the ideal k is bigger, the consulting frequency of a DM should become higher too. Figure 2-1 illustrates our normative predictions of a DM's information buying behavior, in terms of her ideal consulting frequency, in each of the four experimental conditions.

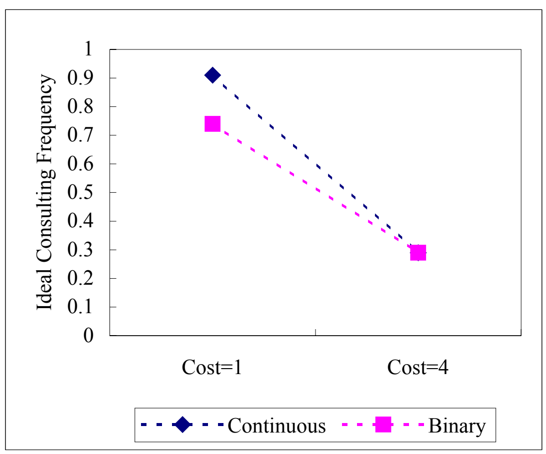

Figure 2-2. Ideal consulting frequency in each experimental condition of Experiment I, inferred from our model-simulations.

Method

Participants

Twelve University of Florida students (7 female) participated in this experiment. All participants had normal or corrected-to-normal visual acuity. Participants were paid

$4 per hour as the base rate plus an incentive bonus that was based on the accuracy of their performance. In addition, in each session of the experimental trials, the person who had earned most points would get an additional reward of $20. By adopting such a payment system, we wanted to motivate our participants to be not only actively involved in the study but also strongly performance-driven.

Apparatus and Stimuli

All stimuli and experiment-related materials were generated and presented in computers, using a program written in Delphi 5.0. Responses from the participants were made through standard computer mice and recorded by our Delphi program. An example of the decision event: a visual stimulus that could be either a "Signal" or "Noise", is shown in Figure 2-2.

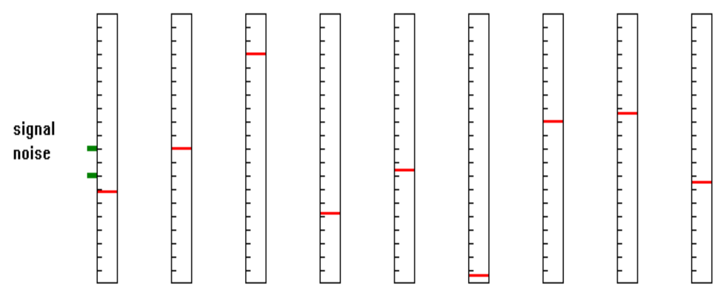

Figure 2-2. An example of the displayed visual stimulus as the decision event.

There were nine individual elements (or gauges) in the stimulus. Each of them consisted of two parallel vertical lines with tick marks on the left line, dividing the gauge into 20 intervals. A horizontal line across each gauge indicated the "value" of that gauge, which could range from 0 to 10 units in height. Those values were determined by sampling from one of two normal distributions N(μ, σ). In the Signal distribution, μ_1=5, and in the Noise distribution, μ_0=4. And the two distributions had a common standard deviation, σ=3.0. On each trial, the nine elements' values were independently drawn from the same distribution. Therefore, they all represented the same kind of decision event: a Signal or a Noise. And the probability of each decision event's occurrence on a certain trial was 0.50, randomly determined by our program.

Based on previous experiments using similar graphical materials (Montgomery & Sorkin, 1996; Luan et al., 2004), the expected performance of a human participant in such a detection task would be around d'=0.75 after training, which is the same as the d'_{DM} value used in our model-simulations. There would be some individual differences across participants, as it would be for any decision tasks. However, because it was a perceptual task in nature, participants' performance should be influenced less by their background, knowledge, etc.. Therefore, we believed that after extensive training, the variance of our participants in d'_{DM} would be small and close to the expected value, d'_{DM}=0.75.

After the presentation of the decision event, participants were asked to report a likelihood rating of the Signal occurrence in a mouse-operated slider. The rating could range from 0: "Very unlikely to be a Signal" (equal to "Very likely to be a Noise") to 100: "Very likely to be a Signal", with 50: "Really not sure what it is" in the middle. This L(x) rating was supposed to be the external representation of a participant's likelihood ratio calculation λ(x), and carry the same mathematical meaning of a posteriori probability estimate, P(s/x).

An advisor's information regarding the same decision event would be available for a participant to consult after she made her initial estimate. It could be either a binary decision or a continuous rating in a certain experimental condition; and was generated by the algorithms described in the experiment-simulations part of the "Simulations" session.

The Cost for accessing the advisor's information would be either 1 or 4 points, varied in different experimental conditions. And the other properties of the advisor's information were the same as listed in Table 2-2. Moreover, to facilitate a participant to understand the crucial properties of the advisor, information of the past performance of the advisor was given to her. And this information was different under the binary and continuous information Mode, as shown in Figure 2-3.

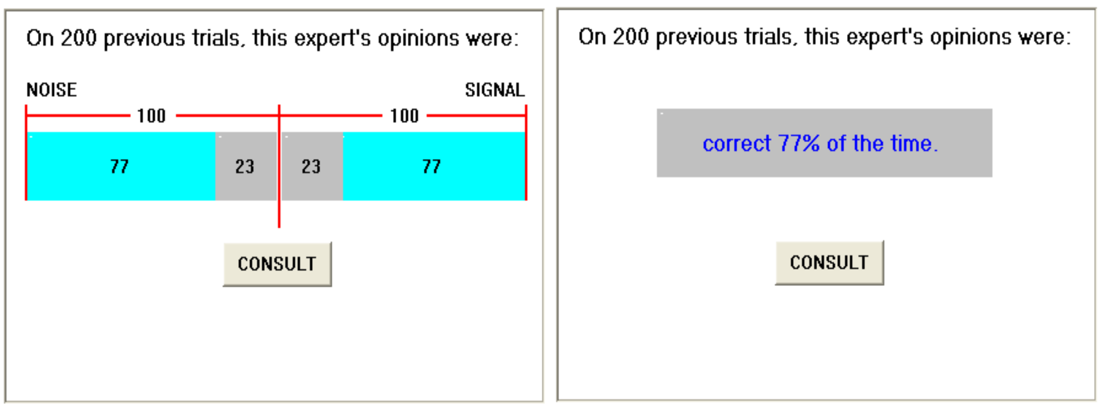

Figure 2-3. Past performance information of the advisor in the binary (left) and continuous (right) Mode conditions in Experiment I.

In the binary Mode, the four rectangular areas, in the order from left to right, represented the number of Hits, Misses, False Alarms, and Correct Rejections resulted from the advisor's binary decisions on the past 200 trials. From the sizes of the areas, a participants could infer such information: 1), the four probabilities: P(s/S), P(n/S), P(s/N), and P(n/N). Thus, it was possible for them to calculate the two likelihood ratios, $\lambda(x/s)$ and $\lambda(x/n)$, which would be useful for information integration. 2), the overall percent correct (PC). It was equal to the total size of the two areas in the leftmost and the rightmost. 3), the decision tendency of the advisor. If more "Noise" decisions were made than "Signal" decisions (the total size of the two areas in the left was bigger than the total size of the two right areas), we referred that advisor to be more "Conservative"; If

"Signal" decisions were made at the equal number of "Noise" decisions, the advisor was referred as being "Neutral"; and finally if more "Signal" decisions were made than "Noise" decisions, we referred that advisor to be more "Liberal".

The decision tendency of the advisor was directly related to its response criterion, β; and the following relations were true: if $\beta>1$, "Conservative"; if $\beta=1$, "Neutral"; and if $\beta<1$, "Liberal". In this experiment, because β_{DM} was equal to 1, the advisor was always being "Neutral". The size of each area was also numbered to enable the information more accurately perceived by the participants.

Past performance information in the continuous Mode was displayed much simpler: just a percent correct number. It was made this way because the advisor's response criterion or decision tendency was not relevant in this case. The number "77%" was the percent correct that was resulted from the decisions made by an advisor with $d'_A=1.50$ and $\beta_A=1$. And the sizes of the four areas in the binary Mode display were determined by the same reason too.

Furthermore, to make sure that the information conveyed in the past performance displays was properly understood by the participants, instructions from the experimenter were given before the onset of the experimental sessions to teach them how to read the information from the displays. A participant could choose at her will to buy the advisor's information or not after viewing its past performance information. Following either action, she would be asked to make a final "yes-no" decision to indicate whether she thought the event was a Signal. For each correct decision she made, she would be rewarded with 10 points; and she would be penalized 10 points for each incorrect decision (see Table 2-2). Feedbacks were given at the end of each trial, including the correctness of the participant's final decision, the point(s) she paid to consult, the total points she earned on that particular trial, and the cumulative points she earned so far in current block of trials.

Procedure

Participants received intensive training (at least 1,200 trials) prior to the experimental sessions. During training, there was no advice available for them to consult. Instead, participants were immediately asked to make a "yes-no" decision after reporting a likelihood rating. The purposes of those training trials were to get our participants familiar with the signal detection task and to stabilize their performance without consulting (the d'_{DM}).

A within-subject design was adopted in this experiment. Therefore, each participant went through all four (2 Mode x 2 Cost) experimental conditions; and the orders of conditions were counterbalanced. In each condition, there were two different sessions with 200 trials in each. The first one (the learning session) was designed for the participants to learn how to buy the advice in a proper way in order to maximize their total points. They were instructed to try any strategy they thought might work for this purpose. In this session, each token point was converted to 0.5 cent (in US dollar) as the performance bonus.

In the second session (the competition session), not only was the value of each token point raised to 1 cent, but also was a participant told that she was in a competition. She would compete with other people who were in the same condition as hers to get a $20 prize, rewarded to whoever came out with the most points. By those manipulations, we hoped that participants were motivated to perform at their best in this session. (For the purposes of this experiment, we decided to only report results from the competition session, although it would be interesting to compare results from both sessions.)

Participants were required to finish 400 trials in each experiment period, which could range from 1.5 to 2 hours. They were encouraged to have breaks during the time, and there was no time pressure for them to finish the trials quickly. Plus training, all participants completed at least 6 experimental periods in this experiment.

Results and Discussion

Manipulation Checks

Two important decision variables need checking for their obtained values in our experiment. The first one is d'_{DM}, the accuracy measure of participants' initial estimates. In all four experiment conditions (200 trials in each), the averaged d'_{DM} value was found no significantly different from the expected value: d'_{DM}=0.75. The second one is d'_{A}, accuracy of the simulated advisor's information, including trials both when it had been consulted by the participants and not consulted. Similarly, in all conditions, the averaged d'_{A} value was found no significantly different from the expected value: d'_{A}=1.50. (For simplicity's sake, I chose not to list those results in this dissertation.) Therefore, the normative buying behavior prescribed by our model-simulations should be valid to be compared to the obtained results.

Consulting Frequency

"Consulting frequency" refers to the frequency or percentage of times that a participant consulted the advisor in a certain experimental condition. Figure 2-4 illustrates the obtained average consulting frequency under each condition, and their corresponding ideal consulting frequencies by our model-simulations.

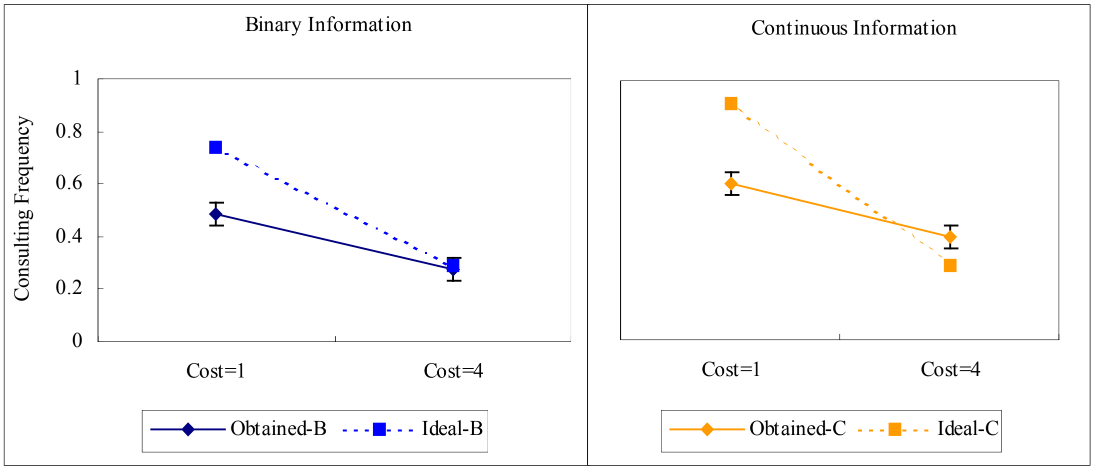

Figure 2-4. Ideal and obtained consulting frequencies (averaged) under each experimental condition in Experiment I.

Through comparisons across the two graphs it can be seen that: Participants consulted the advisor more often when the continuous information from the advisor was available, no matter whether the information cost was low or high ($F(1,11) = 67.18$, $p < 0.01$). And they also consulted the advisor more often when Cost was low, under both the continuous and binary Mode conditions ($F(1,11) = 6.84$, $p = 0.024$). Although those two main effects were significant, the interaction of the two manipulated variables, Cost and Mode, was not ($F(1,11) = 0.01$, $p = 0.99$). This pattern of results was somewhat inconsistent with our model's prescriptions, where the interaction effect was supposed to be significant.

Comparing the obtained to the ideal consulting frequencies, we can see that: When Cost was low, participants consulted the advisor less frequently than they should have under both the binary and continuous Mode conditions, $t(11) > 4.00$, $p < 0.01$. When Cost was high, they consulted as frequently as they should in the binary Mode condition, $t(11) = 0.36$, $p = 0.36$; but consulted more often than they should in the continuous Mode condition, $t(11) = 2.34$, $p < 0.05$. These results indicate that participants "under-consulted" the advisor even when its information was cheap to get; however, when it was expensive, they either "over-consulted" or consulted in a proper portion of times, depending on what type of information was given from the advisor.

Consulting Frequency by Initial Rating

According to our model, a DM should consult the advisor whenever her initial likelihood rating falls in the consulting range, $[L(x_c\text{-}k), L(x_c+k)]$; and not consult at all if her rating is out of that range. If we plot a DM's consulting frequency as a function of the ratings, an "all-or-none" type of graph, with a frequency of 1 in the middle and 0 on the sides, should be observed.

Because participants' initial ratings were in the range from 0 to 100 in the experiment, to make it simpler to draw the graph, we cast the initial ratings into 10 categories. Ratings ranging from 0 to 10 were put in category "1", 11-20 in category "2", and so on. We calculated participants' consulting frequencies in every rating category, and then averaged them across all participants. After those calculations, we plotted the averaged consulting frequency as a function of the initial rating categories, as illustrated in Figure 2-5.

A clear result from this figure is that participants did not consult the advisor in an "all-or-none" fashion. In all conditions, they consulted more frequently when their initial ratings were closer to the middle categories. It means that the stronger they thought their observations were, the less likely they would purchase information from the advisor. They consulted "stochastically" rather than "deterministically". Therefore, there are no clear buying criteria we can identify by which participants made their consulting decisions[1].

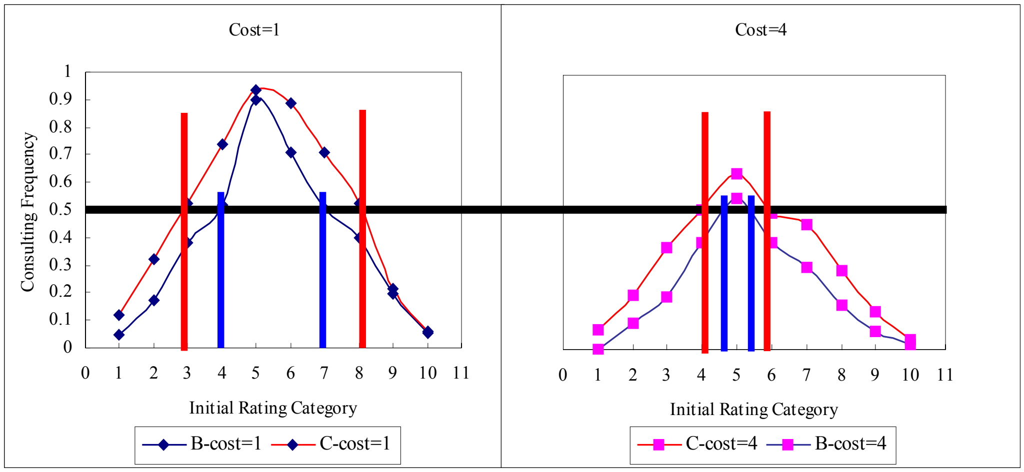

Figure 2-5. Average consulting frequency as a function of the initial rating categories under each experimental condition in Experiment I.

To clarify this issue, we used a method that may enable us to get a rough estimate of those criteria: Using the frequency of 0.50 as a cutoff value to distinguish the "more-likely-to-buy" zone and the "less-likely-to-buy" zone, as can be seen by the line across the middle of Figure 2-5. The crossing points between this line and the consulting frequency functions can tell us at which initial rating category that the participants became more likely to consult the advisor; and at which category that they again became more unlikely to consult (from left to right). By this method, we can see from Figure 2-5 that: The consulting zone was bigger under the continuous Mode; and it was also bigger when the Cost was lower. Those results are consistent with the main effects found in the overall consulting frequency analysis.

Decision Performance

We used d' as the main measurement of participants' decision performance inferred from both her initial ratings (Initial) and her final binary decisions (Final) in the four experiment conditions. Because the patterns of those results were very similar in the two

Mode conditions, we chose to just report the results in the binary Mode condition, as shown in Figure 2-6.

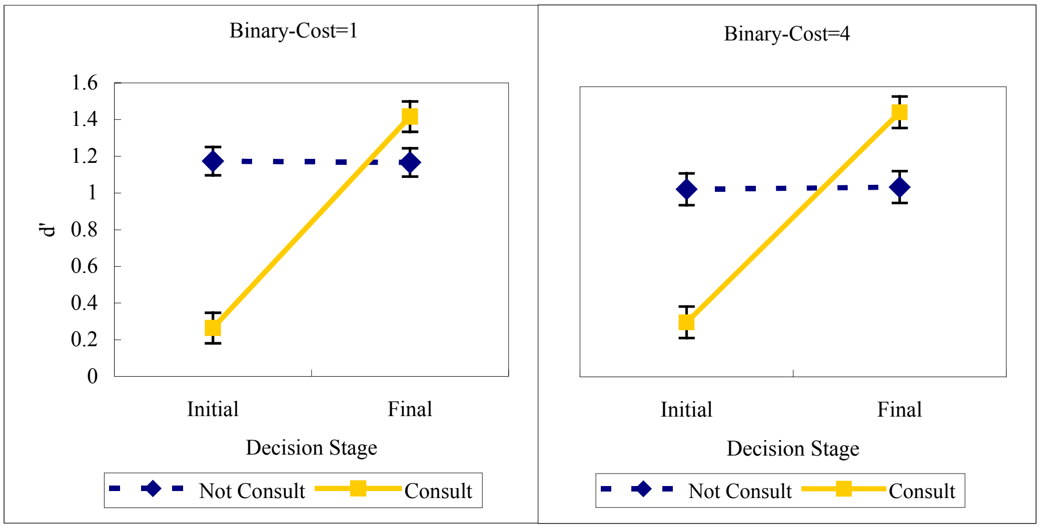

Figure 2-6. Averaged Initial and Final decision performance (in d') in the Binary-Cost=1 and Binary-Cost=4 conditions in Experiment I.

As seen in the figure, participants' Initial performance was much lower when they consulted the advisor than when they did not, $t(11) = 12.99$, $p < 0.01$. However, their Final performance was higher after they consulted the advisor than when they did not, $t(11) = 3.26$, $p < 0.01$. Also, there was no change between the Initial and the Final performance when they did not consult the expert, $t(11) = 1.47$, $p = 0.09$; but a big increment when they did, $t(11) = 14.19$, $p < 0.01$. These results indicate two things: First, participants had a fairly accurate sense of when to consult and when not to: they consulted only when their Initial performance was poor; and second, while consulting the advisor cost them, their Final performance significantly improved because of the integration of the advisor's opinions.

Another important performance measure is the average points participants earned on a trial. We thought that it would be meaningful to compare the earning a participant might have gotten if she had decided not to consult the advisor (an imaginary earning), with the earning she actually got by consulting the advisor (an obtained earning), on trials when she did consult the advisor. The imaginary earning was calculated based on a participant's initial ratings, and the obtained earning was calculated from her final decisions. Again, the patterns of this result were very similar in the two Mode conditions; and we again chose to just report the results in the binary Mode condition, as shown in Figure 2-7.

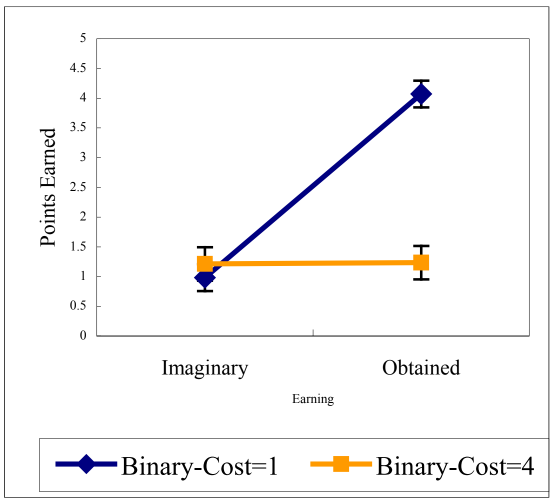

Figure 2-7. Averaged Imaginary and Obtained earning (in points) in the Binary-Cost=1 and Binary-Cost=4 conditions.

It can be seen that: When the information cost was low (Cost=1), consulting the advisor helped participants earn much more points than had she decided not to consult, $t(11) = 11.92$, $p < 0.01$. However, when the information cost was high (Cost=4), consulting actually did not help participants earn more points, $t(11) = 0.25$, $p = 0.403$. Therefore, although consulting in high Cost situation improved participants' accuracy, it actually did little to increase their earnings. This finding is important because it indicates that accuracy improvement sometimes can be meaningless if we do not consider the Cost factor in the advice-taking process.

*Note 1: One may question that the reason why the stochastic consulting was found in the results is because we took the average consulting frequency of our participants. It is possible that every participant used a deterministic consulting strategy but had different consulting cutoff points (buying criteria) in their initial ratings. And if that were the case, averaging the consulting frequencies from all participants would result in a deceptive pattern of stochastic consulting. However, after examining each participant's data, this was found not to be the case. Therefore, averaging did not change the actual pattern of the results.

CHAPTER 3
EXPERIMENT II

Advice-taking should become more complicated in this experiment, when two advisors with different decision properties would be available for a DM to consult. The DM would have the freedom to consult none, one, or both of them at her will; and she could buy one piece of information at a time (using the “optional” stopping rule). Our main interest in this experiment was to further probe the accuracy-cost tradeoff process in advice-taking. By making the two advisors have different expertise (d’) as well as different information costs (Cost) to access their advice, we wanted to determine from a DM’s consulting preference which aspect: accuracy or cost, would be weighed more heavily during the tradeoff.

The properties of the two advisors were designed in this way: One of them (A) would have a higher d’ than the other one (B). Therefore, A’s advice would be more accurate than B’s; and consulting A would help a DM make more accurate final decisions than consulting B. However, A’s information would be more expensive to get than B’s. Thus, if a DM needed help from external sources, she had to make a decision: “Which advisor should I consult first: the more accurate one A, or the cheaper one B?” To make this design more meaningful, the selection of those two advisors should meet this criterion: Consulting either advisor would result in the same expected value for the DM, regardless of which one the DM consulted first. Reflected in our model, similar EV curves as a function of k should be found by consulting either advisor first.

If this criterion is met, it should not matter to the DM which advisor she consults, because they would be equally useful to her in the sense of maximizing her expected total payoff (EV). However, if she shows a preference to consult one particular advisor, we can infer which variable weighs more heavily in her tradeoffs: the expertise of the advisors (the representation of “accuracy”), or the information cost (the representation of “cost”). For example, if the DM consults the high-d’-high-Cost one (A) more often, it will indicate that a higher final decision accuracy is what she thinks more important to maximize her EV. In contrary, if she tends to consult the low-d’-low-Cost one (B) more often, it will indicate that reducing consulting cost is what she thinks more important to achieve her goal of EV maximization. Therefore, our design would enable us to find which side of the tradeoff is more emphasized in a DM’s mind: accuracy or cost.

After trying different d’s and Costs by running similar model-simulations as described in Chapter 2, the following specific property values were found to fit for our design purposes: d'_A=1.50 and $Cost_A$=3 for advisor A, and d'_B=0.85 and $Cost_B$=1 for advisor B. The other important decision properties that were controlled in the simulations are listed in Table 3-1.

Table 3-1. Values of the controlled variables in Experiment II.

d'_{DM}	β_{DM}	β_A	β_B	$\rho_{DM\text{-}A\text{-}B}$	Mode (A & B)	D-payoffs (points)	
0.75	1	1	1	0	Continuous	Hit: 10	False Alarm: -10
						Miss: -10	Correct Rejection: -10

Because a DM could consult none, one, or both advisors in this experiment, we discuss the EV curves in two conditions differently: a, When the DM chooses to consult at most one advisor (0 or 1); and b, when she chooses to consult up to two advisors (0, 1, or 2). Figure 3-1 shows the EV curves as the function of k resulted from consulting either advisor in condition a. It can be seen that the selection of those two advisors meets our design criterion closely: Although the tail part (k>1.3) of consulting B results in higher EVs than the tail part of consulting A, their ideal ks and MaxEVs are almost the same.

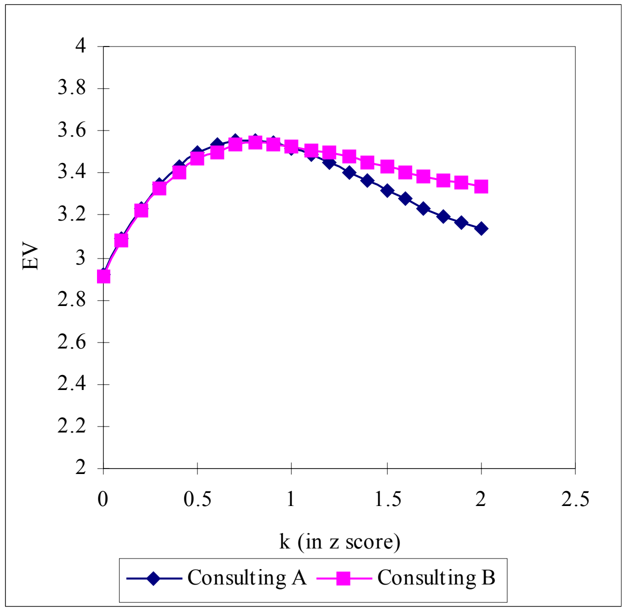

Figure 3-3. EV curves as the function of k by consulting either advisor A or B in the condition where a DM can consult up to one advisor in Experiment II.

Because a DM sometimes needs to integrate two instead of one advisors' information, optimal advice-taking can be more complicated in condition b. Based on the same principles of our model, the following explains what the normative treatments are in this situation: As in the simple case of only one advisor available, a DM needs to select two criteria: $\lambda(xc_0\text{-}k)$ and $\lambda(xc_0+k)$, to decide whether to buy advice (see Figure 1-4). If x, her observation, falls in between the two criteria, she should buy information from one of the two advisors first, says A. After integrating information from A, if her updated criterion λ(DM-A) is still falling between $\lambda(xc_0\text{-}k)$ and $\lambda(xc_0+k)$, she should continue buying information from advisor B; otherwise, she should stop buying and make a final decision. After integrating B's information, if her updated criterion λ(DM-A-B) is larger than her decision criterion β_{DM}, she should make a Signal final decision; otherwise, a Noise final decision.

This algorithm was used in our model-simulations for condition b. And because there are two possible consulting sequences for the DM: consulting A first then B (A-B) or consulting B first then A (B-A), we compared the results from both sequences, as shown in Figure 2-2. It can be seen that both of them result in similar ideal ks and MaxEVs. Therefore, it should not matter for the DM to consult which advisor first. Thus, the selected decision properties of advisor A and B meet our design criterion in both condition a and b.

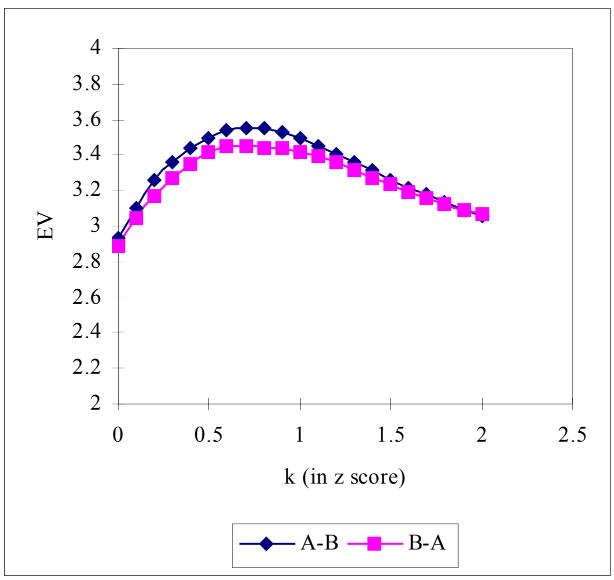

Figure 3-2. EV curves as the function of k in consulting sequence A-B and B-A in the condition where a DM can consult up to two advisors in Experiment II.

Method

Participants

Six University of Florida students (4 female) who had participated in Experiment I participated in this experiment. They were paid $4 per hour as the base rate plus an incentive bonus that was based on the accuracy of their performance. In addition, the same “competition prize” of $20 as used in Experiment I would be rewarded to the person who had earned most points in an experimental session.

Apparatus, Stimuli, and Procedure

The same kind of visual stimuli as used in Experiment I were displayed as the decision events. And the same facilities, computers and mice, were used to conduct the experiment and record participants’ responses. The procedures of an experiment trial were very similar to Experiment I’s only with an addition of the second advisor. Because both advisors (or “Experts” as displayed) would give the participants their continuous information only, percent corrects were displayed as the information of their past performance, as shown in Figure 3-3. In addition, the Costs of their information were also displayed.

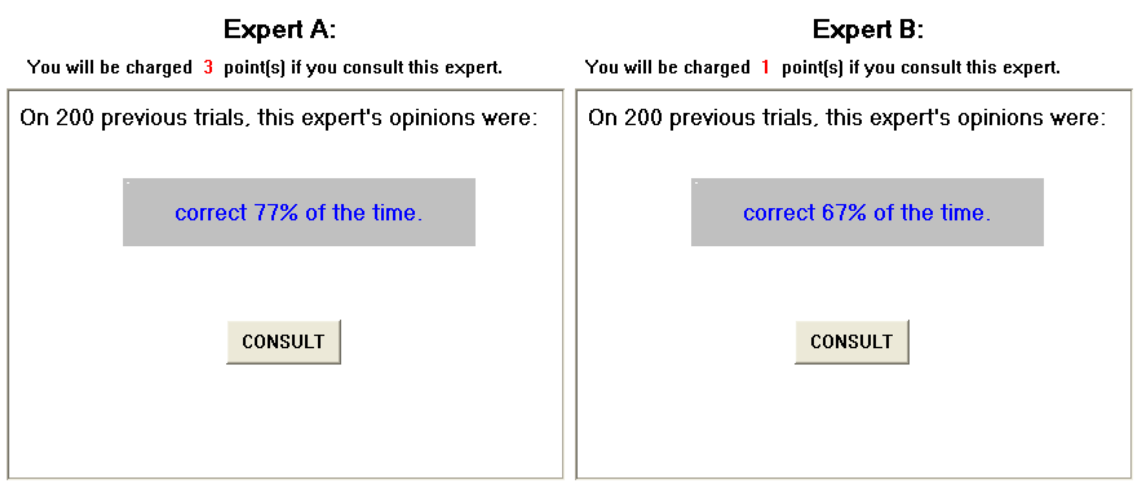

Figure 3-3. Past performance information of the advisors and their consulting Costs in Experiment II.

Because all participants had experience with the basic signal detection task in Experiment I, there was no need for them to do the training without any advisor available to consult. However, they did receive another kind of training that was similar to the Experiment I task. In these training trials (400 in total), they would be given the opportunity to either not consult or consult one available advisor. And in half of the times, that advisor had properties as what advisor A had; and in the other half, its properties were the same as advisor B's. The purpose of those "single-advisor" trainings was to get our participants familiar with the two advisors, so that they could be prepared to choose which one they would like to consult first in the experiment sessions.

There were two conditions in this experiment. In the first condition, which we called "optional-consulting", participants could consult any number of advisors they wanted by using the optional stopping rule. This condition was the major condition of Experiment II; and each participant had completed 600 trials in this condition (200 for learning and 400 for competition). At the end of the experiment and after participants finished all trials in the optional-consulting condition, they would be asked to do another 200 "forced-consulting" trials. In this condition, participants were required to choose one of the two advisors to consult, no matter whether they wanted it or not. The reason why we wanted to add this condition was to see whether a consulting pattern found in the optional-consulting could be carried out in this forced-choice-like situation too. If a participant showed consistent consulting preference in both conditions, it would prove that her consulting preference was reliable.

Results and Discussion

Manipulation Checks

The accuracy of participants' initial estimates (d'_{DM}), advisor A's information (d'_{A}), and advisor B's information (d'_{B}) were checked for their obtained values in the experiment setting. They were all found not significantly different from the expected values: d'_{DM}=0.75, d'_{A}=1.50, and d'_{B}=0.85. And again, for simplicity's seek, I chose not to list those results in this dissertation.

Consulting Frequency

Figure 3-4 illustrates the percentage of times that each participant consulted 0 (p_0), 1 (p_1), or 2 (p_2) advisors in the optional-consulting condition. As we can see, participants did not consult the advisors often, mean (p_1+p_2)=0.292, s.d.=0.15. Compared to the ideal total consulting frequency, which would be 0.545 inferred from Figure 3-2 by converting ideal k to cumulative probability, they significantly under-consulted the advisors, $t(5) = 4.32$, $p < 0.01$. And after participants consulted one advisor, most of they barely consulted the second one. In fact, two participants (1 & 6) did not consult a second advisors at all, p_2=0.

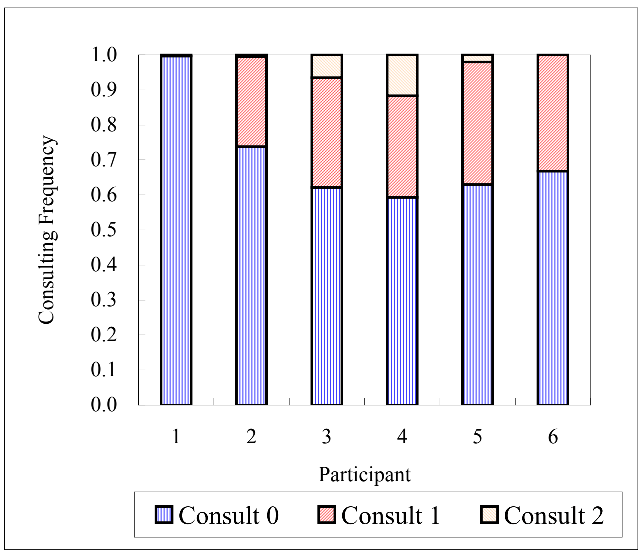

Figure 3-4. Consulting frequency of each participant in the optional-consulting condition of Experiment II.

Participants' consulting frequencies were so low in this experiment that they were even lower than their own consulting frequencies in the Continuous-Cost=4 condition of Experiment I, $t(5) = 1.72$, $p = 0.07$ (marginally significant). Why is it so? One reason may be that: Unlike Experiment I, participants had to make a decision which advisor they would like to consult first in this experiment. Because neither advisor had an advantage over the other in helping the participants to maximize their total payoffs, this consulting decision may construe a “high-conflict” decision situation for them. According to Tversky and Shafir (1992), this may promote a decision-maker to stay with her default decision choice (the decision based on her initial estimate in this study), and flight away from the high-conflict situation. As a result, participants were reluctant to consult either of the advisors. A second possible reason is related to the participants’ consulting preference, and we will present that result first and get back to the discussion later.

Consulting Preference

Figure 3-5 illustrates the averaged percentages of times that participants consulted advisor A first or advisor B first in both the optional- and forced- consulting conditions; and A was the “high-d’-high-Cost” advisor and B was the “low-d’-low-Cost” one. In both conditions, their frequencies of consulting B first were significantly larger than 0.50, $t(5) > 9.00$, $p < 0.01$. Therefore, it can be inferred that participants preferred to look at B, the low-d’-low-Cost advisor’s information first when they decided to consult. And their preference to this advisor was consistent no matter whether they had the freedom to consult any number of the two advisors, or they were forced to consult one of them.

According to our design, because both advisors were equally useful to a participant, she should be indifferent at which one she would choose to consult first. If a consistent preference pattern was observed, it can tell us which variable was weighed more heavily in her mind: the expertise of the advisors (accuracy) or their consulting costs (cost). It turned out from the results that our participants were more sensitive to the cost of the advisors rather than their expertise. When one of the advisors could offer them more accurate opinions, its higher consulting cost “scared” the participants away to purchase the other advisor’s less accurate but cheaper information.

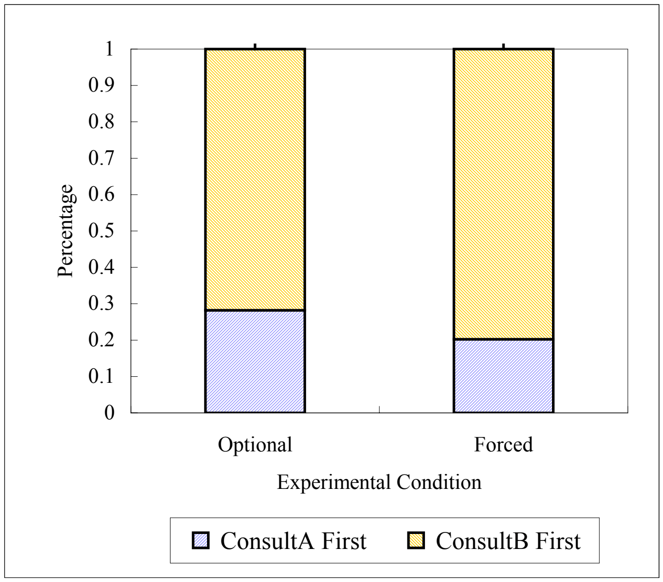

Figure 3-5. Participants' consulting preference in both the optional and forced consulting conditions in Experiment II.

Following the discussion left in the previous session, we think that the second reason why participants consulted too little could be a dynamic process between consulting preference and consulting frequency. Because participants preferred the low cost advisor (B), the information they got from B could improve their final decision accuracy only a little because of its low expertise. When participants found out (through feedback) this drawback of consulting B, they might begin to question whether it was worthy of paying for its opinions at all, even though the cost was low. Because participants were sensitive to Cost and the preferred advisor's information was not reliable, they might choose to rely on their own decision estimates rather than on consulting to make a final decision. As a result, they were reluctant to consult either of the advisors much.

A DM's sensitivity to Cost is not an uncommon finding in the judgment and decision-making literature. According to Kahneman and Tversky's prospect theory

(1982), people are more risk-averse to a certain amount of “loss” than being risk-seeking to the same amount of “gain”; and more risk-averse to a sure loss than being risk-seeking to a probability gain. Consulting a high-d’-high-cost advisor, is like a “big sure-loss, high probability-gain” gamble prospect; while consulting the low-d’-low-cost advisor is like a “small sure-loss, low probability-gain” prospect. Because the design of our experiment made the expected values of those two prospects equal, therefore, by the principles of the prospect theory, people would prefer the “small sure-loss, low probability-gain” prospect to the “big sure-loss, high probability-gain” prospect, which was consistent with what we found in this experiment’s results.

CHAPTER 4
EXPERIMENT III

In this experiment, two advisors would be available for the participants to consult; and their advice would both appear in the binary information Mode. The decision properties of those advisors would be exactly the same except for their decision criteria. While one of them (A) would have a more "Conservative" criterion ($\beta_A>1$), the other one (B) would have a more "Liberal" criterion ($\beta_B<1$). Therefore, if consulted, advisor A would tend to give a participant more "Noise" advice than "Signal"; while B would tend to do the contrary. Those tendencies of their advice would be reflected in the displayed past performance information of the two advisors; and a participant would be asked either "optional-consult" any number of the advisors, or "forced-consult" one of them (like the two conditions in Experiment II). Several consulting strategies were tested in our model-simulations to find which one would be the best for a DM to use in such task situations. The main purposes of this experiment were to see which strategy or strategies our participants would actually adopt to consult the two advisors; and whether they were able to use the best strategy as our simulations prescribed.

Two of the tested strategies are what we call "response-based" strategies, because which advisor a DM would consult first will depend on what her initial estimate is. In one of them, the "Low-C-High-L" strategy, a DM will consult the Conservative advisor first if her initial estimate is lower than 50 (therefore indicating a "Noise" initial decision). If her estimate is higher than 50 (therefore indicating a "Signal" initial decision), she will consult the Liberal one first. Because a DM will always consult the advisor whose opinions are more likely to agree with her own, this strategy can be deemed as a "confirming" strategy. The other strategy, "Low-L-High-C", is exactly the opposite of the "Low-C-High-L" one. Therefore, it can be deemed as a "disconfirming" strategy.

By the definitions of those two strategies, it is not hard to see the similarity between the Low-C-High-L strategy and the so-called "confirmation bias" phenomenon in information searching. It has been found in many studies that people prefer seeking the kind of information that they know will support their own thoughts, as opposed to the conflicting information (e.g., Klayman & Ha, 1987; Jonas, Schulz-Hardt, Frey, & Thelen, 2001). This confirmatory searching strategy is deemed as a bias because it often violates the normative information gathering rules and has been found in lots of cases to result in suboptimal decision performance (e.g. Nemeth & Rogers, 1996). Nonetheless, it is widely adopted by human beings to guild their pre-decisional information acquisition. In this experiment, we were curious how well the Low-C-High-L strategy, a confirming strategy in nature, would do to help a DM. And we were also curious how well the Low-L-High-C, the disconfirming strategy, would do too. By comparing the performance resulted by using those two strategies, we would be able to find whether "confirmation bias" is actually a good strategy or indeed a bad one in advice-taking.

The other two tested strategies are called "preference-based" strategies, because in those strategies which advisor a DM consults first will not depend on her initial estimates but her persistent preference to one particular advisor. One of them is the "Prefer-C" strategy, in which a DM will consult the Conservative advisor first in most of the times. And the other one is called the "Prefer-L" strategy, in which a DM will favor the Liberal advisor to consult first. Two variations of those strategies: "Always-C" and "Always-L", in which a DM will always consult the same particular advisor first, would be tested in our model-simulations. Although they are the simplified versions of the "Prefer-C" and "Prefer-L" strategies, the results can be representative.

In this experiment, advisor A and B's decision criterion was: $\beta_A=2$ and $\beta_B=0.5$, respectively. And the values of other important controlled task variables are listed in Table 4-1. By adopting a criterion at $\beta_A=2$, the probability that advisor A would make a Noise response on a given trial is around 0.64; while the probability that A would make a Signal response is around 0.36. And the reverse results will occur for advisor B's binary responses. Using those task values, we ran similar model-simulations like in the previous two experiments in both the optional- and forced- consulting conditions.

Table 4-1. Values of the controlled task variables in Experiment III.

d'_{DM}	β_{DM}	$d'_{A\&B}$	$Cost_{A\&B}$	$\rho_{DM\text{-}A\text{-}B}$	Mode (A & B)	D-payoffs (points)	
0.75	1	1.50	3	0	Binary	Hit: 10	False Alarm: -10
						Miss: -10	Correct Rejection: -10

In the optional-consulting condition, simulations in two sub-conditions were run separately: a, When the DM chooses to consult at most one advisor (0 or 1); and b, when she chooses to consult up to two advisors (0, 1, or 2). Figure 4-1 and 4-2 show the EV curves as the function of k resulted from the four tested consulting strategies in both conditions. It can be seen that very similar patterns of results occur in both conditions: The confirming Low-C-High-L strategy has the best performance in terms of the MaxEV; the disconfirming Low-L-High-C strategy has the worst; and the performance of the two "preference-based" strategies fall in the between. Moreover, this pattern of results will become more obvious if a DM consults more frequently (when "k" gets bigger).

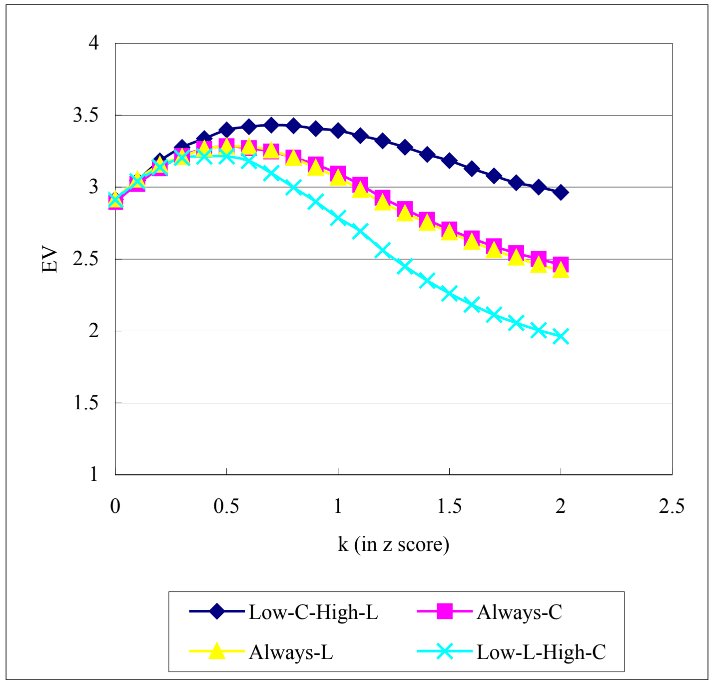

Figure 4-4. EV curves as the function of k resulted from the four tested consulting strategies in the condition where a DM can only consult up to one advisor.

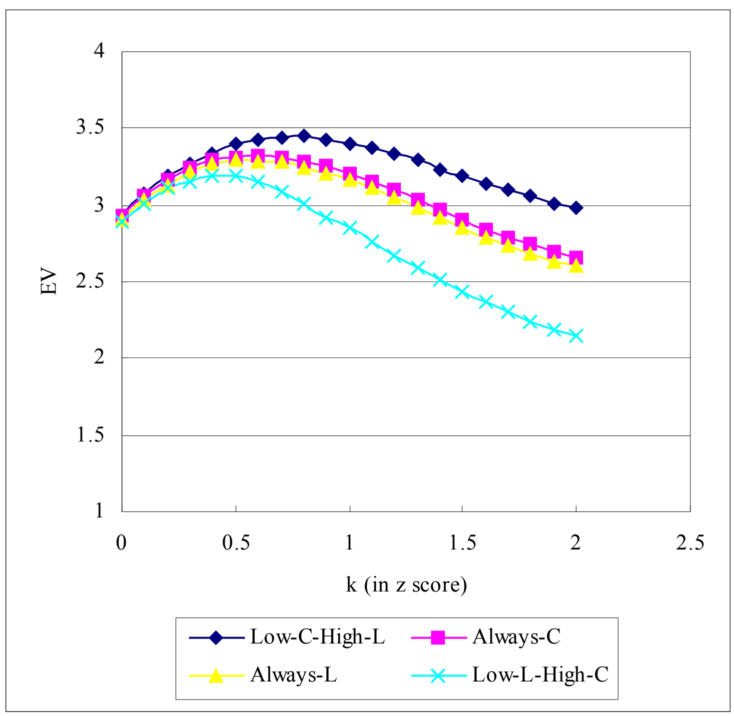

Figure 4-2. EV curves as the function of k resulted from the four tested consulting strategies in the condition where a DM can consult up to two advisors.

We also run simulations in the forced-consulting condition, where a DM will be asked to mandatorily select one advisor to consult on every trial. Because a DM needs not to set any buying criteria in this condition, each tested consulting strategy will result only a single EV, as shown in Figure 4-3. The results are consistent with the results from the two optional-consulting conditions: Low-C-High-L is the best, followed by Always-C and Always-L with almost the same performance; and Low-L-High-C is the worst.

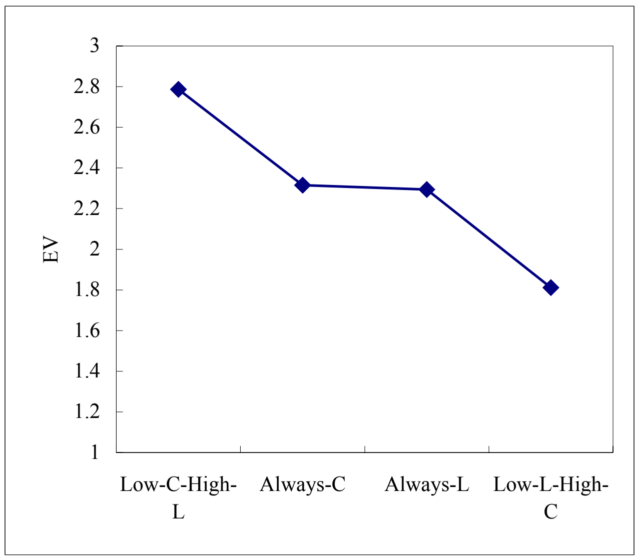

Figure 4-3. EV resulted from each of the four tested consulting strategies in the forced-consulting condition.

Those simulations results may seem to be surprising, because seeking confirmatory information has long been viewed as a faulty bias in human information processing. Not only is it found to be the best advice-taking strategy under present task environments, but also is the opposite of this consulting strategy found to be the worst. To understand better why this pattern of results occurs, we should run more simulations using different values of the crucial task variables. By doing this, we will be able to test how robust this pattern of results is, and define the boundary conditions when it holds and when it does not. (This work is still in progress.)

We speculate right now that there may be three reasons why a confirming strategy works so well in our simulations. First, the main argument why the confirmation bias results in suboptimal decision performance is that confirming sources often bring redundant information only. In many cases, their information is nothing but repeated facts or perspectives that a DM has already known and used to form the basis of her initial decision. Meanwhile, the information from the disconfirming sources can be very useful, because they may remind the DM some important facts and offer insightful perspectives that she is not aware of or has ignored about the decision event. This "diversity of information" can be represented by the information correlations between the DM and advisors (ρ) in a SDT task: The more are an advisor's observations correlated with the DM's, the more possible that its information will be confirming to the DM's, and the less useful will its information be for the DM (see Sorkin & Dai, 1994; Luan et al., 2004). Because all sources' information is set to be independent in our simulations ($\rho_{DM\text{-}A\text{-}B}$=0), this is not the reason why the two advisors would give confirmatory advice to the DM. Therefore, the disadvantage that the confirming information sources might have in other studies does not exist in present task environment; and that may be one reason why adopting a confirming strategy works well here.

The second reason is also related to the fact that all sources are independent in our simulations. If two sources are independent and their opinions happen to agree with each other, the conditional probability that the agreed opinion is correct can be very high (see Luan, Sorkin, & Itzkowitz, 2003). This "agreeing effect" is certainly moderated by the accuracy of the two independent sources, as well as how severely their information is biased (which can be represented by their decision criteria β). It may be true that the setup of our simulations happens to enable a DM to take some advantage of this effect by always consulting the advisor who is more likely to agree with her. To test this hypothesis, we need to vary the information correlations between the DM and the advisors, their decision accuracy, and the values of their decision criteria.

Finally, there may be another advantage by using a confirming strategy. To illustrate, assume that a DM has made a Noise initial decision; but because her observation was so weak, she decides to consult the Conservative advisor A. Based on her experience, A is expected to respond with a Noise binary decision to the DM. If A actually responds a Noise, the DM can take advantage of the agreeing effect and make a Noise final decision. However, if A responds a Signal this time, how should a DM do? She may speculate why A responds differently this time. A highly possible explanation is that A just had a strong observation, and that observation was strong enough to force A to change its usual responses. Considering the fact that her own observation was weak, a reasonable way for the DM to do is to follow A's advice and choose Signal as her final decision. Thus, a DM can take advantage of this unusual response from A and make a more likely correct final decision. This positive effect is what we call the "unorthodox information effect". And it indicates how knowing an advisor's decision tendency can work for a DM's advantage.

Running more simulations with a variety of manipulated variables is our future task to further understand how and when a confirming strategy can be such a good strategy in advice-taking. However, from our current simulations results we can at least draw the conclusion that a searching strategy based on a confirmation-bias-like rationale can sometimes be the best one available for a DM. Therefore, it may be not irrational but even very rational for people to seek confirming evidence. In this experiment, by putting our participants in the same decision environments as our simulations had, we wanted to know whether they could choose this strategy as their strategy-to-use.

Method

Participants

Eight University of Florida students (5 female) who had participated in Experiment I participated in this experiment. They were paid $4 per hour as the base rate plus an incentive bonus that was based on the accuracy of their performance. In addition, the same "competition prize" of $20 used in previous experiments would be rewarded to the person who had earned most points in an experimental session.

Apparatus, Stimuli, and Procedure

The same kind of visual stimuli were displayed as the decision events (refer to Figure 2-2). And the same facilities, computers and mice, were used to conduct the experiment and record participants' responses. The procedures of an experiment trial were identical to Experiment II's. Because both advisors (or "Experts" as displayed) would give the participants only their binary information, rectangular bars and vertical lines were displayed as the information of their past performance, as shown in Figure 4-4. In addition, the Costs of their information were also displayed. From the display, a participant could easily tell the decision tendencies of those two advisors; and identify A as the "Conservative" advisor while B as the "Liberal" one.

Because all participants had experience with the basic signal detection task in Experiment I, there was no need for them to do the training without any advisor available to consult. And there were no those "single-advisor" trainings either, because all needed information about those two advisors would be given in the experimental sessions; and we did not want to prime the participants with the thoughts that one particular advisor should be preferably consulted.

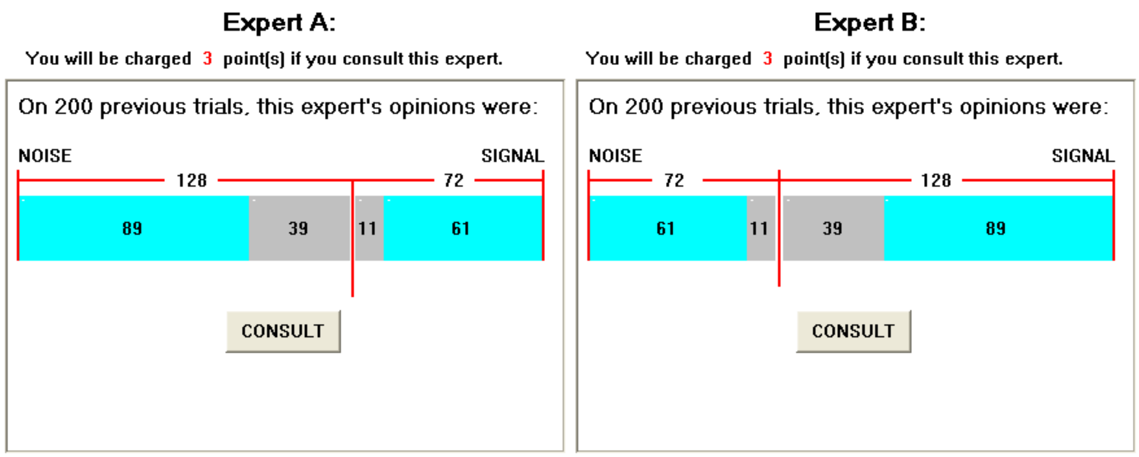

Figure 4-4. Past performance information of the two advisors and the Costs for accessing each of them in Experiment III.

There were two conditions in this experiment. In the optional-consulting condition, participants could consult any number of advisors they wanted by using the optional stopping rule. And each participant had completed 600 trials in this condition (200 for learning and 400 for competition). At the end of the experiment and after participants finished all trials in the optional-consulting condition, they would be asked to do another 200 "forced-consulting" trials. In this forced-consulting condition, participants were required to choose one of the two advisors to consult on each trial, no matter whether they wanted it or not.

Results and Discussion

Manipulation Checks

The accuracy of participants' initial estimates (d'_{DM}), advisor A's information (d'_A), and advisor B's information (d'_B) were checked for their obtained values in the experiment setting. The decision criteria of advisor A and B were also checked. They were all found not significantly different from the expected values: d'_{DM}=0.75, d'_A=1.50, d'_B=1.50, β_A=2, and β_B=0.5. And again, for simplicity's seek, I chose not to report those results in this dissertation.

Consulting Strategy

To identify the consulting strategy or strategies our participants used in the task, we first classified their initial estimates into two categories: Low, with a rating less than 50; and High, with a rating larger than 50. Then, in the condition when they consulted at least one advisor, we calculated their percentages of consulting A (the Conservative advisor) first, and consulting B (the Liberal advisor) first under these two rating categories. Table 4-2 gives an example of such calculations in the optional-consulting condition.

Table 4-2. Identification of each participant's consulting strategy in the optional-consulting condition in Experiment III.

Participant	Initial Estimate				Strategy
	Low (<50)		High (>50)		
	Consult A 1st	Consult B 1st	Consult A 1st	Consult B 1st	
1	0.020	0.980	0.000	1.000	Prefer-L
2	0.887	0.113	0.691	0.309	Prefer-C
3	0.706	0.294	0.300	0.700	Low-C-High-L
4	0.925	0.075	0.265	0.735	Low-C-High-L
5	0.980	0.020	0.033	0.967	Low-C-High-L
6	0.935	0.065	0.099	0.901	Low-C-High-L
7	0.916	0.084	0.075	0.925	Low-C-High-L
8	0.659	0.341	0.171	0.829	Low-C-High-L

If a participant consulted advisor A more often when her initial estimate was low but consulted advisor B more often when her estimate was high, her strategy can be identified as a "Low-C-High-L" strategy. A participant would be identified to adopt a "Low-L-High-C" strategy if the reverse consulting pattern was observed. However, if she consistently preferred consulting one particular advisor (A or B) no matter what her initial estimate was, her strategy could be identified as either “Prefer-L” or “Prefer-C”, depending on which advisor she preferred. By this method, we tried to identify each participant’s consulting strategy in both the optional- and forced- consulting conditions. And the numbers of participants who used each of the four strategies in each condition are listed in Figure 4-5.

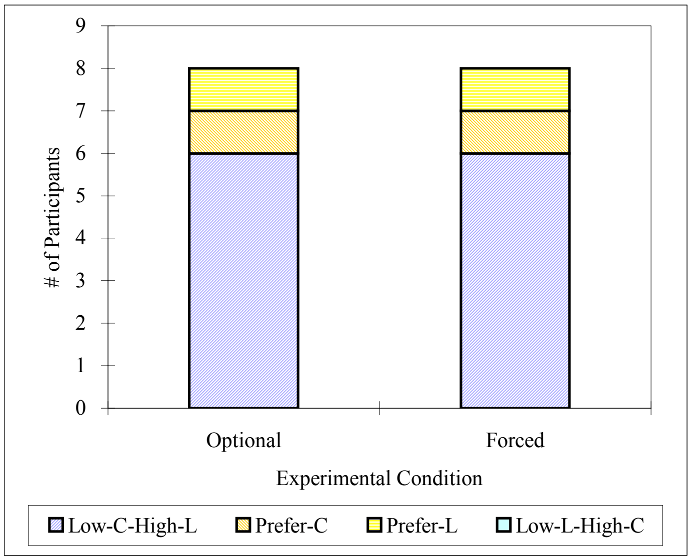

Figure 4-5. Number of participants who used each consulting strategy in the two experimental conditions of Experiment III. Note that no participants used the Low-L-High-C in either of the conditions.

It is clear that the “Low-C-High-L” strategy was the dominant strategy used by our participants. According to our simulation results, this strategy should be the strategy by using which a participant could maximize her total payoff under both the optional- and forced- consulting conditions. Therefore, most of our participants were able to find the best strategy to consult the advisors. Moreover, our simulations also showed that the reverse of this strategy, the “Low-L-High-C” strategy, should be the worst one. As we can see from the results, none of the participants used that strategy in either condition of the experiment. Furthermore, after checking each participant’s data, all participants were found to use one same strategy consistently throughout the two conditions, including the two who used “preference-based” strategies.

Consulting Frequency

Figure 4-6 illustrates the percentage of times that each participant consulted 0 (p_0), 1 (p_1), or 2 (p_2) advisors in the optional-consulting condition. As we can see, the total percentage that the participants consulted one and both of the advisors was fairly high, mean (p_1+p_2)= 0.501 and s.d.= 0.209. However, after participants consulted one advisor, they barely consulted the second one. In fact, four participants (1, 2, 4 & 8) did not consult a second advisor at all, p_2=0.

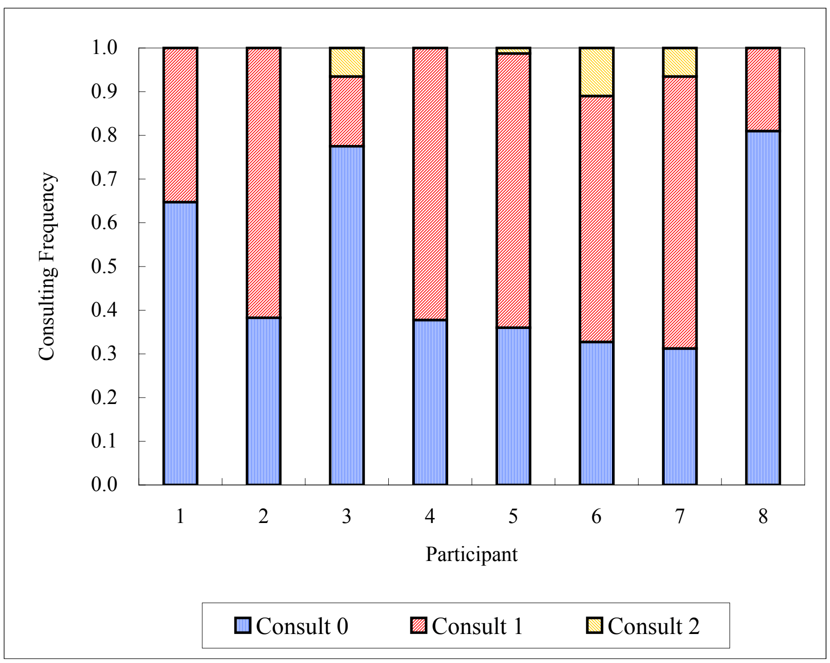

Figure 4-6. Consulting frequencies of each participant in the optional-consulting condition of Experiment III.

From Figure 4-2 in the simulation results, we can see that the ideal k by using the Low-C-High-L strategy was around k=0.7 in *z* score. After converting it to a cumulative probability, the ideal consulting frequency for a participant who adopted that strategy would be 0.486. The averaged total consulting frequency (p_1+p_2) from the six participants who used the Low-C-High-L strategy in the experiment was 0.506, which was not different from the ideal, $t(5) = 0.211$, $p > 0.05$. This result indicates that most of our participants not only were able to use the best consulting strategy, but also they managed to purchase advisors' information a proper percentage of times. It seems that they were not overly sensitive to the Costs of the advisors as were those participants in Experiment II. The primary reason, we think, is that the Costs were set to be equal for the two advisors, and the advisors had equal expertise. This might have helped simplify the accuracy-cost tradeoff process and enable participants to be more focused on finding the best consulting strategy rather than on balancing the improved accuracy and the cost of consulting each advisor.

CHAPTER 5
GENERAL DISCUSSION

In this study, we examined some issues in an empirically important but theoretically still developing area in decision-making: advice-taking. We first presented a model developed from the signal detection theory (SDT) that can give normative prescriptions of how people should acquire and use advice in a defined task environment. Unlike the normative models in the information purchasing area (e.g., Edwards, 1965; Connolly, 1988), our SDT model emphasizes the role of the advice-takers, and treats their own decision estimates as a determining factor about whether to consult external sources for more information. Different from existing research in the advice-taking area (e.g., Harvey & Fischer, 1997; Yaniv, 2004), in our model the cost for accessing advice is treated as a natural feature, and hence a crucial task variable, of an advice-taking task. Moreover, our model regards the pre-decisional information acquisition and the in-decisional information integration as an integrated process in advice-taking and defines normative treatments in both sub-procedures. Therefore, compared to previous studies in related areas, our SDT model is not only a new model to study advice-taking, but also a more thorough one. We hope that research using this model as theoretical guidance can help shed more light on our understanding of human beings' advice-taking behavior.

In the rest of the study, we conducted three laboratory experiments to probe how human participants took advice in various decision situations. Before each experiment, we prescribed how participants should respond in the experimental conditions based on results from computer simulations, with specific values of the task variables as the input.

Through comparisons between the obtained and the simulations results, we were able to see whether our participants could behave normatively; and if not, how much they deviated from the optimal. Furthermore, those laboratory studies would also help us identify certain patterns of people's advice-taking behavior, if there is any.

In Experiment I, it was found that people were sensitive to the changes of the advice-related variables and could adjust their consulting frequencies accordingly. In one aspect, when the cost of accessing the advisor's information increased, participants purchased less information. This result is both intuitively understandable and consistent with the model's prescriptions. If other important properties of the advice such as its information diagnosticity remain unchanged or unimproved, it will be unworthy of still consulting the advisor at a high frequency when its advice becomes more expensive.

In the other aspect, when the content of the advice changed from continuous ratings to binary opinions, our participants consulted the advisor less frequently. Apparently, they thought that a piece of binary advice was not as useful as a continuous one. Hence, they were not willing to purchase both kinds of advice at the same price. Why is it so? From our simulations, we can see that binary advice is inherited to be less diagnostic than continuous advice. One may say that this is the case because of the specific algorithms we used to generate these two kinds of advice. However, we would argue that this is a general fact for most decision studies involving expert estimates.

Imagine, for example, that we want to hear advice from some meteorologists about tomorrow's weather. They can either tell us whether it will rain or not (Binary); or tell us the chance of raining in a probability scale (Continuous). A reasonable assumption is that they first make some analyses from the available weather data and form an estimate how possible it will rain tomorrow. If the probability is higher than a preset "raining criterion", say 0.50, they will then predict that it will rain; otherwise, not rain. (Note that the raining criterion is not necessarily to be 0.50.) However, this will make a 0.99 raining probability no different from a 0.55 one in a binary forecast. And it will certainly affect the actions we will take to prepare for the next day's weather (e.g., taking an umbrella or not). Misses and false alarms are expected to occur more following the binary forecasts, because there is just not enough information for one to make a more proper response.

Another drawback of the binary advice is that it is more difficult for people to use, especially in how to use it optimally. To use binary advice in the ideal way, it requires a DM first to accurately estimate the four probabilities (Hit, Miss, False Alarm, and Correct Rejection) resulted from an advisor's previous decision record. This was made easy in our experiment as all those probabilities were displayed in the advisor's past performance panel. The difficult part is that a DM must convert those probabilities into two likelihood ratios (or furthermore two posteriori probabilities) in order to integrate the binary advice with her own estimate. We doubt that a layperson knows how to finish these calculations correctly, because they are both methodologically and computationally complex. It is highly possible that the value of the binary advice is reduced in a DM's mind because of the cognitive difficulty it brings in integration.

Compared to binary advice, continuous advice not only contains more information but also is easier for a DM to use. Therefore, it may become more attractive to shop for. However, when considering consulting cost in the process, continuous advice may not always be the preferred one. When the consulting cost is high, according to our model a DM should buy advice only when she has a very weak observation of the decision event, because a decision made based on such an observation would be highly possible to be an incorrect one. After a DM obtains the advice in this situation, the probability that she will conform to the advisor's opinion and give up her own should be high, regardless whether it is binary or continuous. Therefore, it will make no difference how much information the advice contains and how difficult it is to be used. The result will be the same: a binary decision that matches what the advice indicates. Our model prescribes that people should buy the binary advice equally frequently as the continuous advice when the Cost is high, because the probability that a DM makes weak observations is independent with the content of the advice.

Apparently, our participants did not quite understand this subtle interaction effect between the content and the Cost of the advice. They constantly discounted the binary advice regardless of whether the Cost was low or high. Because of the possible cognitive drawbacks of the binary advice, we think that this pattern of behavior is understandable, although not totally rational.

We also found in Experiment I that our participants generally had a rational sense of when to consult and when not to. They consulted more often when their initial estimates were low and less often when they were high (and a low expressed initial estimate often led to a less accurate decision). This "leverage" effect of an advice-taker's own decision estimate on her subsequent advice-taking behavior is consistent with our model's prescriptions. However, different from our model, we found that participants made their consulting decisions "stochastically" rather than "deterministically". It seems that they just could not make up their mind whether a particular initial estimate should definitely lead to a consulting action or not. As a result, they sometimes consulted the advisor even when their estimates were very high; and did not consult even when their estimates were very low. We think that this pattern of results reflects participants' perception of the stochastic natures of both the decision events and a decision entity's responses: A strong observation does not always result in a correct decision (it can only lead to a higher likelihood ratio estimate); and there always is a certain probability that an advisor's information can be misleading (none advisor can give fully accurate advice all the time).

As discussed throughout this paper, buying advice can be seen as an accuracy-cost tradeoff process. A direct question about this process is: Which aspect is deemed as more important for a DM to achieve her ultimate decision goal of maximizing the total expected payoff (EV)? In Experiment I, when the Cost was low, participants were "stingy" at purchasing advice. They consulted the advisor less frequently than they should have, a sign indicating that "cost" was weighed more heavily than "accuracy". If there were a consistent pattern, they would have bought advice also less frequently than they ideally should when the Cost was high. However, this was not found to be the case. In the high Cost conditions, participants either over-consulted the advisor when the advice was continuous ratings (a sign of weighing "accuracy" more), or consulted properly when the advice was binary opinions. Therefore, there was no distinguishable tradeoff pattern we could identify in Experiment I.

To further probe this tradeoff process, we conducted Experiment II with such setup: a "high-d'-high-Cost" advisor was available to consult together with a "low-d'-low-Cost" one. By carefully selecting the specific values of the d' and Cost of those advisors, participants should be indifferent about which advisor they would consult, because they would be equally useful to them in achieving their goal of EV maximization. Therefore, if participants showed persistent preference to a particular advisor, we can infer which feature of the advisors attracted them more: the high d' or the low Cost. And we can further infer which aspect was weighed more heavily in the tradeoff: accuracy or cost. This design constructed a direct contrast between accuracy and cost. By kind of forcing a participant to make a choice decision between the two advisors, we hoped that her true preference would be revealed.

In both the optional and forced consulting conditions of the experiment, it was found that our participants preferred consulting the low-d'-Low-lost advisor. Therefore, we draw the conclusion that "cost" was being weighed more heavily in their mind when they tried to balance accuracy with cost. This result can be explained by Kahneman and Tversky's prospect theory (1982), as discussed in Chapter 3. Basically, it says that people are more risk-aversive to a choice option with "big sure loss although high probability gain" than an option with "low probability gain but small sure loss". Therefore, they tend to select the later option when both of them are available in a choice set.

Another major result in Experiment II was that the overall consulting frequency was found to be very low in the optional consulting condition. We speculated that the difficulty of deciding which advisor should be consulted first was one reason why this occurred, because neither of the advisors had an obvious advantage over the other. To avoid this aversive choice dilemma (Tversky & Shafir, 1992) or to reduce the regret that might be caused by selecting one of them to consult but not the other (Zeelenberg, 1999), sticking to the default decision option: their own decision estimates, might seem to be an easy way out for the participants.

From another point of view, this choice difficulty might bring an additional cost to the task: the cost of “thinking” (Russo & Dosher, 1983). Because this kind of cost is hard to measure and be quantified, we did not include it in our formal SDT model. However, it does exist in lots of advice-taking situations; and its magnitude can vary from task to task, as can be seen through task comparisons between Experiment I and II. And it seems that it adds to the “cost” side of the tradeoff and reduces people’s motivation of buying more information from the advisors. To systematically probe the effects of such cost on advice-taking and to find plausible ways to incorporate it into our model, more studies need to be conducted in future.

Compared to Experiment II, the accuracy-cost tradeoff was much easier in Experiment III, because the consulting cost for accessing each of the two advisors was the same and the two advisors had equal expertise (d’). It should be clear to the participants that the task was not mainly about balancing accuracy with cost, but rather finding the best way to consult advisors who gave biased advice. Once a particular consulting strategy was determined, a participant could just apply that strategy whenever she wanted to consult the advisors. And it was found that our participants could consult the advisors in a proper portion of times in the optional consulting condition.

The main purpose of Experiment III was to study how people would take biased advice from the advisors. “Bias” in present study was defined as the decision tendency of a certain advisor and was represented by the criterion measure, β, in SDT. To better communicate the advisors’ biases to the participants, we chose to display their binary opinions only. In that “mode” of information, whether an advisor was being “Neutral”, “Conservative”, or “Liberal” could be discerned by the participants with the aid of the advisor's past performance information and from its responses on experimental trials. Although this was our definition and manipulation of advisors' biased advice, it could have different meanings in other advice-taking situations.

For example, if advice is displayed in the mode of continuous ratings, there can be certain information tendencies of the advice too. For instance, if a rating, say R, is generated by the mechanism used in our study (see Chapter 2), it will be a reflection of the advisor's observation with no alteration. However, such well-calibrated ratings may be rare in a natural decision setting. An advisor may tend to give more "exaggerated" ratings (R+x) for some reasons, like wanting to impose more influence to the DM (Sniezek & Buckley, 1995); or more "humble" ones (R-x) for other reasons, like being cautious not to mislead the DM. Once a certain tendency of the advice is established and can be discerned by a DM, we will say that the advice is "biased" in some way (a more specific classification of decision biases can be seen in Kerr et al., 1996). Because advisors or experts in most circumstances are not immune to biases, studying how people take and use biased advice may be even more realistically meaningful than unbiased advice. Unfortunately, such studies are few and most available ones are more concerned with the "use" (integration) part of advice-taking than the "take" part (e.g., Birnbaum & Stigler, 1979; Kerr et al., 1996; Lim & O'Connor, 1995). Therefore, we consider our Experiment III unique; and hope that findings in this experiment may bring some interesting issues to the study of advice-taking.

The major finding in Experiment III was that most of our participants used a strategy (Low-C-High-L) that is comparable with the so-called "confirmation bias" to take advice from the two available advisors. And this strategy was identified as the best one among four tested strategies by our model-simulations. Because a confirming strategy is often deemed as an inferior and irrational strategy in information processing, our results may seem surprising: Not only can seeking advice from someone who is more likely to agree with one's own proposition be a good strategy, but sometimes it can be the most efficient one. The possible reasons why this holds in the task conditions like Experiment III have been discussed in Chapter 4; and we are still running more simulations to test our hypotheses. However, for whatever reasons, the conclusion we can draw from this experiment is clear: People can be very good at taking biased advice by applying an intelligent strategy to consult the advisors.

One may notice that we did not report much about how our participants actually utilized their purchased advice in this paper, except for some brief performance descriptions and discussions in Experiment I. However, it does not mean that those results are not important. In contrary, we think that to understand the whole process of advice-taking, analyzing how people integrate or use the advice and what their final decision performance are should be just as important as those pre-decisional information acquisition analyses. And by the normative treatments prescribed by our model and techniques developed from SDT (see Sorkin et al., in press), we are well equipped to do so. Unfortunately, limited by the length of this dissertation report and the massive body of such results, they were not made as the foci in this study. We hope that we can report and discuss those results in other manuscripts for the time to come.

Advice-taking is such an important decision practice that it is hard to imagine how people would make decisions, especially the important ones, without seeking opinions from other people or information sources. However, taking advice can be complicated.

Mapping to the task procedure used in the present study, there are at least four processes that may be involved in advice-taking. The first one is “decision evaluation”. Before taking any advice, a DM (the advice-taker) needs to somehow form her own decision estimate based on either her observation of the decision event or other event-related information. Because of the stochastic natures of both the decision event and the DM’s judgment, there is always some uncertainty in this initial estimate; and this uncertainty may drive people to seek external opinions. During this process, factors like the difficulty of the decision task, a DM’s ability to make correct decisions, and her perception of the certainty level, can all contribute to determine whether she needs advice and how much advice she needs.

The second one is “source evaluation”. Before a DM decides to consult an advisor (the advice-giver), she needs to know, much or some, the important characteristics of this advisor, such as its expertise or credibility, whether there is a decision bias in its advice, in which kind of format the advice will be given, etc. Some of this source information may be missing and some of it may need inferring from other relevant information. And in lots of times a DM may form a false impression of the source. However, no matter how the evaluation comes out, it will help the DM to decide or justify whether she needs to consult that advisor or not.

The third one is “accuracy-cost tradeoff”. Considering that there is often a cost associated with a piece of advice, this step is crucial in finalizing which advisor or advisors the DM needs to consult and how many of them. To make proper consulting decisions, the DM needs to carefully balance the gain she expects to receive by consulting the advisor(s) with the cost she will pay. Because all information gathered in the previous two stages, in addition to the cost information, can potentially influence the tradeoff outcome, this process may be very complicated. And the last one of the processes is “information integration”. After a DM finishes consulting an advisor or advisors, she needs to put all the pieces of information together to make a final decision or judgment. Based on the type of the input information, different integration methods may be needed during this process. With the final decision or judgment made, advice-taking ends.

The above descriptions summarize the processes that may be involved in a general advice-taking task. Each of those processes can be individually studied and some have been extensively studied in the broad area of judgment and decision-making. But, under a unified theme of advice-taking, they should not be treated as independent processes with no connections to each other. The product of one process can affect the processing of other processes; and when the feedback information and repeated trials are available, all processes may become contingent on each other. To understand advice-taking better, we think that a broad and Gestalt-like view of the processes needs to be taken rather than a narrow and separated one.

The SDT model developed in this study has tried to take a broader perspective on advice-taking and identify the connections among the processes. However, because of the complex nature of the topic, we can hardly say that our current model is a complete one. Theoretically, there still is plenty of work we need to do in future studies. And because our model is a new one to the area, we were more concerned with what this model could contribute to study advice-taking in general rather than taking a particular topic as the focus of the present study. Therefore, the three experiments we conducted covered more “breadth” than “depth” for the research questions that had interested us. To gain more and deeper understanding of those questions, there is also a lot of empirical work left for us to do for the time to come.

In conclusion, the study reported in this dissertation continued our research line of applying the signal detection theory to the field of judgment and decision-making. The formal SDT model and three laboratory experiments have shown that our approach can be useful in examining research questions in a relatively undeveloped but practically important area: advice-taking. Although this study has covered a broad range of topics in advice-taking, including testing people’s sensitivity to crucial task variables, finding how they traded accuracy with cost, and identifying what kind of strategy they used to take biased advice, we believe that more theoretical and empirical questions centered around advice-taking need and can be studied by our SDT approach in the future. Finally, we hope that our work can attract more research attention from behavioral scientists in psychology and other fields to advice-taking, because it is such a common and critical decision practice in our everyday life.

LIST OF REFERENCES

Anderson, N. H. (1968). Application of a linear-serial model to a personality-impression task using serial presentation. *Journal of Personality and Social Psychology, 2,* 531-539.

Ashton, A. H., & Ashton, R. H. (1985). Aggregating subjective forecasts: Some empirical results. *Management Sciences, 31,* 1499-1508.

Birnbaum, M. H., & Stigler, S. E. (1979). Source credibility in social judgment: Bias, expertise, and the judge's point of view. *Journal of Personality and Social Psychology, 37,* 48-74.

Brehmer, B., & Joyce, J. C. R. (1988). *Human Judgment: The SJT View*. North-Holland: Elsevier Science Publishers B. V..

Budescu, D. V., & Rantilla, A. K. (2000). Confidence in aggregation of expert opinions. *Acta Psychologica, 104,* 371-398.

Clemen, R. T., & Winkler, R. L. (1999). Combining probability distributions from experts in risk analysis. *Risk Analysis, 1999,* 187-203.

Connolly, T. (1977b). Cues, components, and causal structure in laboratory judgment studies. *Educational and Psychological Measurement, 37,* 877-888.

Connolly, T. (1988). Studies of information-purchase processes. In Brehmer, B., & Joyce, J. C. R. (Ed.), *Human Judgment: The SJT View*. (pp. 401-425). North-Holland: Elsevier Science Publishers B. V.

Connolly, T., & Gilani, N. (1982). Information search in judgment tasks: A regression model and some preliminary findings. *Organizational Behavior and Human Performance, 30,* 330-350.

Connolly, T., & Thorn, B. K. (1987). Predecisional information acquisition: Effects of task variables on suboptimal search strategies. *Organizational Behavior and Human Decision Processes, 39,* 397-416.

Connolly, T., & Wholey, D. R. (1988). Information mispurchase in judgment tasks: A task-driven causal mechanism. *Organizational Behavior and Human Decision Processes, 42,* 75-87.

Durlach, N. I., Braida, L. D., & Ito, Y. (1986). Towards a model for discrimination of broadband signals. *Journal of the Acoustical Society of America, 80,* 60-72.

Edwards, W. (1965). Optimal strategies for seeking information. *Journal of Mathematical Psychology, 2,* 312-329.

Einhorn, H. J., & Hogarth, R. M. (1982). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology, 32,* 53-88.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Harvey, N., & Fischer, I. (1997). Taking advice: Accepting help, improving judgment and sharing responsibility. *Organizational Behavior and Human Decision Processes, 70,* 117-133.

Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social Psychology, 80,* 557-571.

Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American, 246,* 160-173.

Kerr, N. L., MacCoun, R. J., & Kramer, G. P. (1996). Bias in judgment: Comparing individuals and groups. *Psychological Review*, *103*, 687-719.

Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94,* 211-228.

Lim, J. S., & O'Connor, M. (1995). Judgmental adjustment of initial forecasts: Its effectiveness and biases. *Journal of Behavioral Decision Making, 8*, 149-168.

Luan, S., Sorkin, R. D., & Itzkowitz, J. (2003). The majority heuristic in information integration. Poster presented at the annual meeting of the Society for Judgment and Decision Making, Vancouver, Canada.

Luan, S., Sorkin, R. D., & Itzkowitz, J. (2004). Weighing information from outside sources: A biased process. *Journal of Behavioral Decision Making, 17*, 95-116.

Macmillan, N., & Creelman, C. (1991). *Detection Theory: A user's guide.* Cambridge, England: Cambridge University Press.

Montgomery, D., & Sorkin, R. D. (1996). Observer sensitivity to element reliability in a multi-element visual display. *Human Factors, 38,* 484-494.

Nemeth, C. J., & Rogers, J. (1996). Dissent and the search for information. *British Journal of Social Psychology, 35,* 67-76.

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

Raiffa, H. (1968). *Decision analysis*. Reading, MA: Addison-Wesley.

Russo, J., & Dosher, B. A. (1983). Strategies for multiattribute binary choice. *Journal of Experimental Psychology: Learning, Memory and Cognition, 9,* 676-696.

Sniezek, J. A., & Buckley, T. (1995). Cueing and cognitive conflict in judge-advisor decision making. *Organizational Behavior and Human Decision Processes, 62,* 159-174.

Sorkin, R. D., & Dai, H. (1994). Signal detection analysis of the ideal group. *Organizational Behavior and Human Decision Processes, 60,* 1-13.

Sorkin, R. D., Hayes, C. J., & West, R. (2001). Signal detection analysis of group decision making. *Psychological Review, 108,* 183-203.

Sorkin, R. D., Luan, S., & Itzkowitz, J. (in press). Group decision and deliberation: A distributed detection process. In Koehler, D., & Harvey, N. (Ed.), *Blackwell Handbook of Judgment and Decision Making*. London: Blackwell.

Stigler, J. G. (1961). The Economy of information. *The Journal of Political Economy, 69,* 213-225.

Tanner, W., & Birdsall, T. (1958). Definition of d’ and η as psychological measures. *Journal of the Acoustical Society of America, 30,* 922-928.

Tversky, A., & Shafir, E. (1992). Choice under conflict: The dynamics of deferred decision. *Psychological Science, 3,* 358-361.

Wald, A. (1947). *Sequential analysis*. New York: Wiley.

Weber, E. U., Baron, J., & Loomes, G. (2001). *Conflict and tradeoffs in decision making*. Cambridge, England: Cambridge University Press.

Wendt, D. (1969). Value of information for decisions. *Journal of Mathematical Psychology, 6,* 430-443.

Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology, 60,* 181-192.

Yaniv, I. (1997). Weighting and trimming: Heuristics for aggregating judgments under uncertainty. *Organizational Behavior and Human Decision Processes, 69,* 237-249.

Yaniv, I. (2004). Receiving other peoples' advice: Influence and benefit. *Organizational Behavior and Human Decision Processes, 93,* 1-13.

Yaniv, I., & Kleinberger, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational Behavior and Human Decision Processes, 83,* 260-281.

Zeelenberg, M. (1999). Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making, 12,* 93-106.

Zimbardo, P. G., & Leippe, M. R. (1991). *The psychology of attitude change and social influence*. Philadelphia: Temple University.

BIOGRAPHICAL SKETCH

Shenghua Luan was born in Helongjiang, China, on February 17, 1977. He attended high school at The First High School Attached to Beijing Normal University in Beijing, China, and graduated in July 1995. He then went to Peking University in Beijing, China in September 1995, and completed his Bachelor of Arts degree in Psychology in July 1999. In August 1999, he began his graduate school study in the Cognitive and Sensory Processes area of the Department of Psychology at the University of Florida; and earned his Master of Science degree in Psychology in May 2002. Now he is expected to complete his Doctor of Philosophy degree in August 2004.

Shenghua's major research interests include: Applied Signal Detection Theory, Group Processes, Information Integration, and Advice Taking and Utilizing. During the course of his graduate studies, he has published one paper in the Journal of Behavioral Decision Making and a book chapter co-authored with his advisor Dr. Robert D. Sorkin and his colleague Mr. Jesse Itzkowitz. He also had over ten conference presentations and a couple of manuscripts in preparation. In April 2003, he was awarded the E. Porter Horne Memorial Scholarship for outstanding graduate student pursuing the study of Sensory Processes, Perception and/or Cognitive Psychology.

Shenghua has accepted a post-doctoral fellowship from the Max Planck Institute in Berlin Germany, and will begin his professional research career from October 2004. He will get married to his love-of-the-world, Ms. Tian Liu, in September 2004; and in his mind this will be the most celebrated thing for him in his whole life.