FORECASTING CORPORATE PERFORMANCE

by

Robert P. Harrington

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Business
with a major in Accounting

APPROVED:

W. E. Leininger, Chairman

L. N. Killough

R. M. Brown

T. W. Bonham

D. J. Johnson

December, 1985

Blacksburg, Virginia

Forecasting Corporate Performance

by

Robert P. Harrington

(ABSTRACT)

For the past twenty years the usefulness of accounting information has been emphasized. In 1966 the American Accounting Association in its State of Basic Accounting Theory asserted that usefulness is the primary purpose of external financial reports. In 1978 the State of Financial Accounting Concepts, No. 1 affirmed the usefulness criterion. "Financial reporting should provide information that is useful to present and potential investors and creditors and other users..."

Information is useful if it facilitates decision making. Moreover, all decisions are future oriented; they are based on a prognosis of future events. The objective of this research, therefore, is to examine some factors that affect the decision maker's ability to use financial information to make good predictions and thereby good decisions.

There are two major purposes of the study. The first is to gain insight into the amount of increase in prediction accuracy that is expected to be achieved when a model replaces the human decision maker in the selection of cues. The second major purpose is to examine the information overload phenomenon to provide research evidence to deter-

mine the point at which additional information may contaminate prediction accuracy.

The research methodology is based on the lens model developed by Eyon Brunswick in 1952. Multiple linear regression equations are used to capture the participants' models, and correlation statistics are used to measure prediction accuracy.

ACKNOWEDGMENTS

I wish to express my sincere gratitude to Professor Wayne E. Leininger for his guidance and assistance throughout this project and to the other members of my committee, Dean Thirdwall W. Bonham, Professors Larry N. Killough, Robert M. Brown, and Dana J. Johnson for their insightful critiques. In addition, I appreciate the support of my wife, who edited this dissertation and shared with me the emotional burden of the doctoral study. This dissertation is dedicated to my parents, and who gave to their children that special something that made this dissertation possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I

INTRODUCTION

Accounting has been defined as, "the process of identifying, measuring, and communicating economic information to permit informed judgments and decisions by users of the information."[1] The role of the accountant, then, is that of a supplier of information. The accountant supplies information to decision makers within an organization for operating, financing, and performance measurement decisions and to external users for credit granting, investing and other types of decisions. In addition to their roles as suppliers of information, accountants themselves are consumers of information and decision makers. Auditors make a wide range of decisions in gathering audit evidence and reporting the results of the audit. Management accountants select information systems and determine the information to be produced for internal and external users.

Usefulness to decision makers is the criterion for the selection of the information to be provided. The American Accounting Association asserted in A Statement of Basic Accounting Theory that usefulness is the primary purpose

---

[1]Committee to Prepare a Statement of Basic Accounting Theory, A Statement of Basic Accounting Theory (Sarasota, Fla.: American Accounting Association, 1966), p. 1.

of external financial reports. The Statement of Financial Accounting Concepts No. 1 states:

> Financial reporting should provide information that is useful to present and potential investors and creditors and other users in assessing the amounts, timing and uncertainty of prospective cash receipts . . .[2]

Information is useful to the extent that it facilitates decision making. Abdel-Khalik states that, "the relevance of a given piece of data is evaluated by its differential effects on decisions."[3]

Prediction is an integral part of decision making. Libby asserts that an information system must be judged on its ability to predict relevant events and that decision making is based in prediction.[4] Ashton presents the predictive ability criterion, "predictive ability is important because predicting is a necessary, and prior condition for decision making."[5] He also states that accounting data should be evaluated in terms of its ability to make

---

[2]Financial Accounting Standards Board, Statement of Financial Accounting Concepts No. 1 (Stamford, Connecticut: FASB, November, 1978), p. viii.

[3]A. Rashad Abdel-Khalik, "The Effect of Aggregation Accounting Reports on the Quality of the Lending Decision: An Empirical Investigation," Supplement to Journal of Accounting Research (1973), p. 105.

[4]Robert Libby, "The Use of Simulated Decision Makers in Information Evaluation," The Accounting Review (July 1975), p. 475.

[5]Robert H. Ashton, "The Predictive Ability Criterion and User Prediction Models," The Accounting Review (October 1974), p. 720.

predictions of future events that are relevant to the decision process--the predictive-ability criterion.[6] "According to this criterion, alternative accounting measures are evaluated in terms of their ability to predict events of interest to decision makers. The measure with the greatest predictive power with respect to a given event is considered to be the 'best' method for the particular purpose.[7] To fulfill their dual roles as both providers of information for decision makers and decision makers, accountants must make an effort to understand both their own decision processes and the decision processes of those who provide the demand for their product, information.

The focus of concern, then, is on the quality of decisions made by both accountants and users of accounting information and methods of improving the decision process. This concern has lead to a facet of accounting research called human information processing (HIP). The purpose of HIP research is to "describe actual decision behavior in accounting contexts, to evaluate its quality, and to suggest remedies for any discovered deficiencies."[8]

---

[6] Ibid.

[7] Committee on Accounting Theory Construction and Verification, "Report of the Committee on Accounting Theory Construction and Verification," Supplement to The Accounting Review (1971), p.61.

[8] Robert Libby, Accounting and Human Information Processing: Theory and Applications (Englewood Cliffs, New Jersey: Prentice-Hall, 1981), p. xi.

1.1 PURPOSE OF THE STUDY

In conformance with the predictive ability criterion, this research examines ways in which predictions, and thereby decisions, can be improved. Libby lists three options for improving decisions, options that are equally applicable for improving predictions:

1. Change the information provided to the decision maker.
2. Educate the decision maker to change the way he processes information.
3. Replace the decision maker with a model.[9]

The third option is the focus of a great deal of research in both psychology and accounting. Much of this research has used the "lens" model, developed by Eyon Brunswich in 1952, to capture the decision maker's model and to measure prediction accuracy.[10] In traditional lens model research, when prediction accuracy is examined, the decision maker is provided with a predetermined set of cues and is requested to predict an event that is known to the researcher.[11] The decision maker's model is captured

---

[9] Ibid., p. 3.

[10] Eyon Brunswick, <u>The Conceptual Framework of Psychology</u> (Chicago: University of Chicago Press, 1952).

[11] Traditional lens model research as it pertains to prediction accuracy is discussed by Robert Ashton. Robert H. Ashton, <u>Human Information Processing in Accounting</u> (Sarasota, Florida: American Accounting Association, 1981).

by a multiple linear regression, and prediction accuracy is measured using correlation statistics. The results of traditional lens model research have been, with few exceptions, consistent: the model outperforms the human decision maker. This research will be replicated in the current study.

In traditional lens model research the decision maker is provided a set of cues. Because the cues are provided, the achieved prediction accuracy is a result of the decision maker's usage of the cues. There is, however, a second major component of prediction accuracy which relates to Libby's first option, information choice. Actual prediction situations require not only the use of cues, but also their selection. The cues can be selected either by the human decision maker or by a mathematical model. According to lens model theory, high prediction accuracy will be achieved when the cues used to make the prediction are highly correlated with the event being predicted.[12] It would therefore be expected that a multiple regression equation would achieve a more optimal cue selection, hence a higher prediction accuracy, than the human decision maker.

Libby discussed the importance of understanding how information is currently being used as compared to optimal methods of information use. ". . . before one can decide

[12] Ibid., p. 30.

that a change is necessary, a baseline is needed to measure the incremental benefits of the change."[13] A major purpose of this research is to gain some insight into the amount of increase in prediction accuracy that is expected to be achieved when a model replaces the human decision maker in the selection of the cues.

There may be a fourth option for improving prediction accuracy; providing the human decision maker with the optimal number of cues. Research has indicated that human decision makers have limitations as to the number of information cues that they can use effectively. A ceiling on prediction accuracy achieved by the human decision maker is sometimes reached after the examination of a small number of cues. The addition of new cues causes prediction accuracy to remain constant or decrease. This phenomenon is called information overload. Slovic and Litchenstein state that "there is a small amount of evidence that increasing the amount of information available increases his [the decision maker's] confidence without increasing the quality of his decisions. . . ."[14] A second major purpose of this research is to examine the information overload phenomenon

---

[13]Libby, Accounting and Human Information Processing, p. 4.

[14]Paul Slovic and Sarah Lichtenstein, "Comparison of Baysian and Regression Approaches to the Study of Information Processing Judgment," Organizational Behavior and Human Performance (1971), p. 687.

to provide some research evidence to determine the point at which additional information may contaminate prediction accuracy.[15]

The context in which information choice and information overload are to be examined is the prediction of return on total assets (ROA) for corporations.[16] Although earnings per share (EPS) has "long been viewed as the single most important indicator of corporate success by the individual shareholder, the financial community, and the business leader,"[17] it has been shown that a change in ROA is a precursor to change in EPS. When the ROA trend is downward, EPS will eventually decline; and when the ROA trend is upward, EPS will eventually rise.[18]

1.2 SIGNIFICANCE

There are two major components of prediction accuracy: the selection of the information cues to be used in making the prediction (information choice), and the use of the selected cues (information use). The essence of human infor-

---

[15]For the purpose of this research, information overload is defined as the point in the prediction process at which the addition of new informatin decreases prediction accuracy.

[16]ROA is defined as net income divided by total assets.

[17]Albert D. Burger and Stewart K. Webster, "The Management Accountant Looks at EPS vs. ROA: Conflict in Measuring Performance," Management Accounting (August 1978), p. 21.

[18]Ibid., p. 20.

mation processing research is to determine the role of the human decision maker in both information choice and information use. Einhorn states that "one must distinguish between the mode of **data collection** and the **combination method** used to deal with the data once collected. Both these factors could involve clinical or statistical methods . . ."[19] A great deal of empirical research evidence has indicated that the statistical combination of cues often generates a higher prediction accuracy than the human combination of the cues.[20]

There has, however, been little research to determine how effectively the human decision maker can perform the data collection and data selection function. Goldberg states the need for research comparing human versus statistical data collection and data selection:

> . . . it now seems safe to assert rather dogmatically that when acceptable criterion information is available, the proper role of the human in the decision-making process is that of a scientist: (a) discovering or identifying new cues which will improve predictive accuracy, and (b) constructing new sorts of systematic procedures for combining predictors in increasingly more optimal ways. The resulting success of the clinician-as-scientist, when he is relegated to such an hypothesis-testing role, obviously must be

---

[19]Hillel J. Einhorn, "Expert Measurement and Mechanical Combination," Organizational Behavior and Human Performance (February 1972), p. 86.

[20]This research is reviewed in Chapter Two.

evaluated empirically.[21]

The significance of this research is that it attempts to provide empirical evidence as to how well human decision makers can perform their roles as scientists in the collection and selection of information to be used in the prediction process.

1.3 METHODOLOGY

The models of the decision makers will be captured using a group of cases in which the decision makers are given cues and asked to predict a future return on assets for a corporation. The Brunswick lens model, with multiple linear regression equations, will be used to measure prediction accuracy and to determine some of the causes of inaccuracy.

First to be examined will be the change in prediction accuracy that is expected to take place when the decision maker is removed from the prediction process and replaced by a mathematical model. The general hypothesis is that prediction accuracy will increase when the man is replaced by the model.

Second, information overload will be examined. It is hypothesized that prediction accuracy will be maximized

---

[21] Lewis R. Goldberg, "Man Versus Model of Man: A Rationale, Plus Some Evidence, for a Method of Improving on Clinical Inferences," Psychological Bulletin (June 1970), p. 423.

after the subject has examined a small number of cues, and a decline in prediction accuracy will then take place as new cues continue to be examined.

An empirical study will be performed using actual financial data from ninety real firms selected from the COMPUSTAT tapes. The ninety cases will be randomly divided into three groups of thirty cases each. Group I cases and Group II cases will be used to examine the impact of information choice and information use on prediction accuracy. The participants will be selected from students enrolled in the Masters of Accountancy and the MBA programs at Virginia Tech.

1.3.1 Information Choice and Information Use

Group I cases will be examined by the participants first. The participants will be given a list of cues and requested to select the six cues they believe to be the most predictive of future ROA. The data for the cues selected will then be provided by the researcher for each of the thirty cases, and the participants will record their predictions of ROA for each case. The participants will then examine the Group II cases in which the participants are provided with six cues that have been selected using multiple linear regression to maximize the multiple correlation coefficient between the cues and the firms' ROAs.

Of prime concern is the measurement of the prediction accuracies of both the participants and multiple regression models created from each of the case sets. The prediction accuracy of the regression models is their multiple correlation coefficents, and the prediction accuracy of the participants is defined as the correlation between the participants' predicted ROAs and the actual ROAs for the firms. Generally, the participants' abilities to select cues (information choice) will be determined by comparing the prediction accuracies of the Group I cases, in which the participants selected their own cues, and the prediction accuracies of the Group II cases, for which the cues were provided. The participants' abilities to effectively use the cues (information use) will be determined by comparing the prediction accuracies of the participants to that of the regression models.

1.3.2 Information Overload

A second factor relating to the participants' abilities to effectively collect and select information to be used for decision making is the number of cues that a human decision maker can effectively integrate without a loss of prediction accuracy, information overload. A second group of participants and the Group III case set will be used in the information overload experiment. Eight cues will be selected using multiple linear regression. The participants will be given a list of the ten cues and requested to select the two cues they believe to be the most predictive of future ROA. The data for the cues selected will then be provided by the researcher for each of the thirty cases, and the participants will record their predictions of the ROA and the level of their confidence in their predictions for each of the cases. After completing the thirty cases, the participants will be instructed to select two additional cues. The data for these cues will be provided, in addition to that previously provided, and the participants will be directed to revise their predictions and their levels of confidence based on the new information. This process will be repeated until all eight cues have been used in the prediction process and the initial predictions and confidence levels have been revised three times.

Prediction accuracy is measured as the correlation between the participants' predictions of ROA and the actual ROA. Information overload will be deemed to have occurred if there is a permanent decrease in prediction accuracy after the selection of the first two cues.

1.4 ORGANIZATION OF THE STUDY

In Chapter Two Brunswick's lens model will be described, along with selected major research studies pertaining to information choice and information use. Additionally, information overload will be presented. The emphasis will be placed on human information research relating to prediction accuracy.

The research questions and the hypotheses to be tested will be presented in Chapter Three. The research methodology, including the selection of the participants, the creation of the cases, a description of the experiments, and the statistical procedures to be employed to analyze the results, will be presented in Chapter Four.

Chapter Five will contain the statistical analyses and results of the experiments; Chapter Six, the results of the research and recommendations for future research.

Chapter 2

LITERATURE REVIEW

Ashton states that Carl Devine is the first accounting researcher to emphasize the behavioral aspects of accounting.[22] Devine stated his concern for the behavioral issues in accounting thus:

> On the balance it seems fair to conclude that accountants seem to have waded through their relationships to the intricate psychological network of human activity with a heavy-handed crudity that is beyond belief. Some degree of crudity may be excused in a new discipline, but failure to recognize that much of what passes as accounting theory is hopelessly entwined with unsupported behavior assumptions is unforgivable.[23]

As a result of concern for behavioral issues, behavioral accounting research became popular in the 1960's and 1970's and included such diverse topics as the impact of budgets and management control systems on employee attitudes and performance, the impact of accounting information on the decision making behavior of both accountants and users of accounting information, human resource accounting, and social responsibility accounting.

Two problems surfaced as a result of the behavioral ac-

---

[22] Robert H. Ashton, Human Information Processing, p. 2.

[23] Carl T. Devine, "Research Methodology and Accounting Theory Formation," The Accounting Review (July 1960), p. 394.

counting research during the 1970's. The first problem recognized was that of quality control. "The quality control problem arose, in part, because of insufficient knowledge of behavioral research methodologies -- by those who attempted to do behavioral research and those who evaluated it."[24] A second problem was "the lack of theoretical frameworks to guide the research efforts . . . the lack of formal models. Without some type(s) of theory(ies) or model(s), each behavioral study stood alone."[25] Because of the lack of theoretical framework, it was difficult for the researcher to formulate hypotheses, identify dependent variables, and determine the extraneous variables that required control.

In the 1970's accounting researchers discovered two paradigms that alleviated the quality control problem and provided a theoretical framework for behavioral accounting research. The subjective expected utility (SEU) paradigm models decision making under uncertainty and is based on the assumption that individuals make choices that maximize their expected utilities. It does not hold that individuals assess probabilities and weigh utilities, but that the maximization of utilities will enable the researcher to predict the choices of individuals in a risky environ-

---

[24]Ashton, Human Information Processing, p. 4.

[25]Ibid.

ment. The SEU paradigm focuses on the use of probabilities in decision making, the measurement of probabilities and the revision of probabilities using Bayne's Theorem.

The lens model, like the SEU paradigm, models decision making under uncertainty. The lens model is based on the construction of linear mathematical equations as representations of an individual's decision or prediction process. There are three major areas of lens model research. The first is policy capturing, the capturing in the form of mathematical models the judgment policy or prediction strategy of the individual. Policy capturing, like SEU, is a "black box" research strategy; it does not attempt to explain the individual's actual mental processes, but to represent the relationships between the inputs and outputs of the decision making process.

The second major area of lens model research is the examination of judgment accuracy and judgment agreement. Judgment accuracy is partitioned into components using the lens model equation so that the determinants of judgment accuracy and judgment agreement can be examined.

The third major area to which lens model research has been applied is multiple cue probability learning (MCPL). The objective of MCPL research is to determine the environmental conditions that affect learning. In the typical MCPL experiment the participants are provided with different types of feedback, and the changes in their judgment accuracy, as a result of the different types of feedback, are measured.

The focus of this research study is on the components of prediction accuracy and the discovery of factors that affect prediction accuracy. Libby states that there are three available options for improving prediction accuracy:

1. Changing the information.
2. Educating the decision maker to change the way he or she processes information.
3. Replacing the decision maker with a model.[26]

Before it can be determined how to best improve decision performance, one must determine how decisions are currently being made and measure current decision performance. Understanding how decisions are currently being made and the flaws in current decision processes provides cues to specific methods of improving decisions.[27] There are two fundamental research questions:

1. How accurate are predictions?
2. How might prediction accuracy be improved?

Discussed in Chapter Two is lens model research pertaining to these two fundamental research questions. The lens model has its roots in cognitive psychology, and research in both the areas of psychology and accounting will be reviewed

---

[26]Robert Libby, Accounting and Human Information Processing, p. 10.

[27]Ibid., p. 4.

in this chapter. The lens model research to be reviewed pertains primarily to judgment accuracy, emphasizing Libby's third option for improving prediction accuracy. First, the Brunswick lens model paradigm and the components of prediction accuracy will be presented.

2.1 THE LENS MODEL

Brunswick's lens model, developed by Eyon Brunswick in 1952, provides the conceptual and methodological framework for this research study.[28] The basic proposition of the lens model is that human decision makers must rely on probabilistic information in making predictions about an uncertain environment. The probabilistic information is in the form of items of information, cues, which are imperfectly related to the event being predicted, the environmental event. The imperfect relationships between the individual cues and the environmental event are measured via correlation coefficients. The quality of the predictions, prediction accuracy, is also measured using correlation coefficients. There are three aspects of the lens model framework; the model of the environment and the individual, the indices of prediction accuracy, and the lens model equation. An understanding of each is required for this study.[29]

---

[28]Eyon Brunswick, The Conceptual Framework of Psychology.

[29]The discussion of the three factors in the framework of the lens model is based primarily on Ashton, Human Information Processing in Accounting, pp. 14 -18.

2.1.1 The Model of the Environment and the Individual

The Brunswick lens model (Figure 1, p. 20) has two focal points; the left side of the model focuses on the environmental event being predicted, and the right side on the individual decision maker's prediction system. A multiple linear regression equation is used in an attempt to select from the environment an optimal set of probabilistic cues so as to achieve the most reliable prediction of the environmental variable. This multiple linear regression equation is the environmental model and takes the form:

$$\hat{Y}_e = b_0 + b_{e1}X_1 + b_{e2}X_2 + \ldots + b_{en}X_n$$

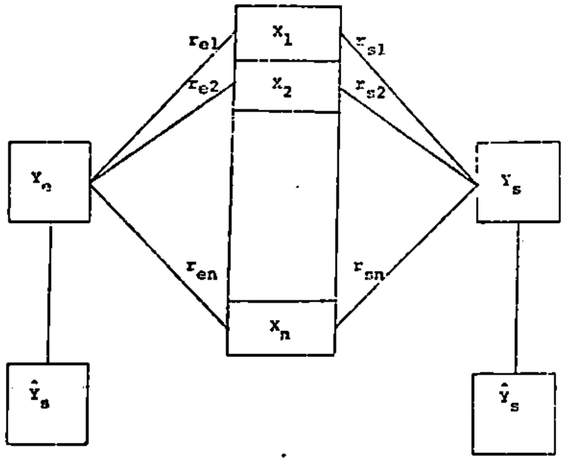
$\hat{Y}_e$ is the predicted value of the environmental event, and the betas indicate the relative weights of the individual cues used in the prediction of Y_e.

The right side of the model focuses on the prediction making process of the decision maker who is using the set of cues to predict the environmental event. The decision maker's prediction model is also represented by a multiple linear regression equation:

$$\hat{Y}_s = b_0 + b_{s1}X_1 + b_{s2}X_2 + \ldots + b_{sn}X_n$$

The betas represent the relative importance attached to each cue by the decision makers in making their predictions of the environmental event.

The lens model methodology, then, is based on the representation of the multivariate relationships existing on each side of the lens as multiple linear regression equa-



Y_e - Distal or criterion variable: the environmental event to be predicted by the decision maker. The distal variable for this study is ROA.

r_{ei} - Validity coefficient: the correlation between the individual cues and the distal variable, ROA.

X_i - The cues: the items of information to be used in predicting the distal variable.

Y_s - The decision maker's prediction of the distal variable.

r_{si} - The utilization coefficient: the extent to which the decision maker uses the individual cues in the prediction of the distal variable.

$\hat{Y}_e$ - The optimal prediction of the distal variable using a multiple regression equation:

$$\hat{Y}_e = b_0 + b_{e1}X_1 + b_{e2}X_2 + \ldots + b_{en}X_n$$

Given cues one through n, the dependent variable is the predicted value of the distal variable. Each beta is an optimal multiple regression beta weight determined by the validity of each cue.

$\hat{Y}_s$ - The optimal prediction of the decision maker's responses using a multiple regression equation:

$$\hat{Y}_s = b_0 + b_{s1}X_1 + b_{s2}X_2 + \ldots + b_{sn}X_x$$

Figure 1

BRUNSWICK'S LENS MODEL

tions. It is important to recognize that the equations do not capture the actual cognitive processes of the decision makers. Instead, ". . . judgment policies captured in this way are intended to represent the relationships among inputs and outputs. They are not intended to explain the 'actual' mode of information processing used to form judgments."[30] While not capturing the cognitive processes of the decision maker, the multiple linear regression equations do provide a basis for examining some factors that affect prediction accuracy.

2.1.2 Indices of Prediction Accuracy

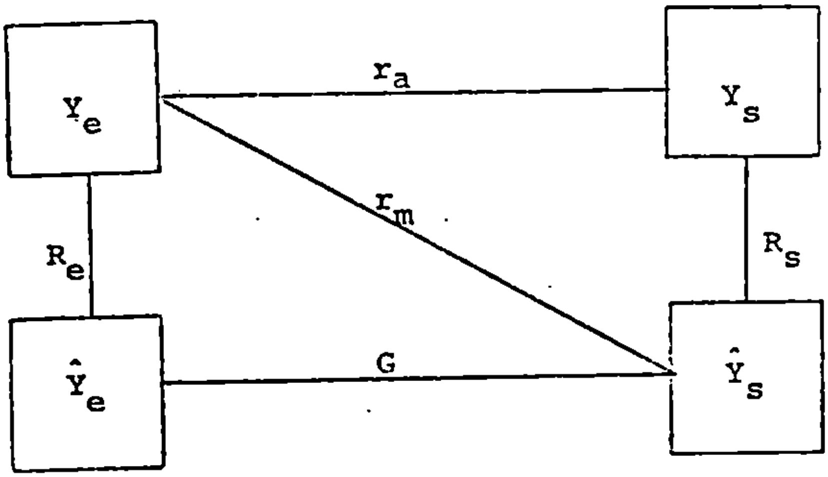
The values of Y_e, Y_s, $\hat{Y}_e$, and $\hat{Y}_s$ are correlated to generate five sets of indices that are employed to measure prediction accuracy and to determine the factors that affect prediction accuracy. The values Y_e, Y_s, $\hat{Y}_e$, and $\hat{Y}_s$ and the indices of prediction accuracy r_a, R_e, G, R_s and r_m are diagramed in Figure 2, p. 22. The indices are interpreted as follows:

r_a - Achievement index: the correlation between the decision maker's prediction (Y_s) and the distal variable (Y_e). $r_a = r_{YsYe}$. r_a measures the decision maker's ability to predict the distal variable.

R_e - Environmental predictability: the multiple correlation coefficient between the distal variable (Y_e)

---

[30] Ibid., p. 13.



Symbol	Name	Definition
r_a	Achievement index	r_{YeYs}
R_e	Environmental predictability	$R_{Ye\hat{Y}s}$
R_s	Consistency index	$R_{Ys\hat{Y}s}$
G	Matching index	$r_{\hat{Y}e\hat{Y}s}$
r_m		$r_{\hat{Y}sYe}$

Figure 2

INDICES OF PREDICTION ACCURACY

and the mathematical prediction of the distal variable (Y_e). $R_e = R_{Ye\hat{Y}e}$. R_e indicates the degree to which the environmental event can be predicted by the cues.

G - Matching index: the correlation between the prediction of the decision maker's response ($\hat{Y}_s$) and the prediction of the distal variable ($\hat{Y}_e$). $G = r_{\hat{Y}s\hat{Y}e}$. G measures the decision maker's ability to weight the cues in an optimal manner.

R_s - Consistency index: the multiple correlation coefficient between the decision maker's response (Y_s) and the decision maker's model's prediction of that response (Y_s). $R_s = R_{Ys\hat{Y}s}$. R_s indicates the extent to which the decision maker's response can be predicted by the cues.

r_m - The correlation between distal variable (Y_e) and the prediction of the decision maker's response (Y_s). $r_m = r_{Ye\hat{Y}s}$. r_m indicates the degree to which the multiple linear regression equation that represents the decision maker's model can predict the criterion values.

2.1.3 The Lens Model Equation

The lens model equation explains prediction accuracy, r_a, in terms of the indices that form the components of prediction accuracy. The equation is stated as follows:

$$r_a = R_e\ G\ R_s + C\sqrt{1-R_e^{\ 2}\ 1-R_s^{\ 2}}$$

R_e sets the upper limit on prediction accuracy. G measures the extent of the decision maker's knowledge of the task; the correct weighting of the cues. G will become unity when the regression coefficients of the environmental model and the decision maker's model are proportional. When G is unity, the decision maker is using the cues in the optimal manner. When G is less than unity, a systematic error is created by the imperfect weighting of the cues by the decision maker; and prediction accuracy is reduced.

R_s measures the extent to which the decision makers consistently apply their own models. When R_s is less than unity, prediction accuracy has been reduced as a result of random error created by the inconsistent application of task knowledge.

C is the correlation between the residual variances of the environmental model and the decision maker's multiple regression model. C represents the decision maker's use of the cues in a nonlinear fashion. Although the lens model equation differentiates between the linear and nonlinear components of prediction accuracy, the nonlinear component being to the right of the plus sign, research has found that the nonlinear component contributes little or nothing to prediction accuracy.[31] The lens model equation

---

[31]Goldberg reviewed a number of studies relating to the use of cues in a nonlinear fashion by decision makers and found that the simple linear model accounted for 80 to 100% of the variance of judgment accuracy in the studies review-

can therefore often be reduced to its linear component:

$$r_a = R_e \, G \, R_s$$

The relationships among the correlation statistics r_a, R_e, G, and R_s are diagramed in Figure 2, p. 23.

2.2 COMPONENTS OF PREDICTION ACCURACY

All approaches to the prediction process involve the human decision maker and/or mathematical models. Libby provides an example, the graduate business school admissions decision. The decision to accept or reject a student applicant is based on a prediction of the applicant's future success in the job market. Success is measured in dollars. The prediction of future success is made by evaluating the cues provided by the applicant; GMAT scores, grade point average, quality of undergraduate school attended, recommendations, participation in extracurricular activities, and answers to subjective questions.

There are several potential approaches to predicting the individual applicant's future success in the job market. An admissions officer could make subjective estimates based on the cues. Individual members of a committee could make subjective estimates and combine the estimates. The admissions officer could provide subjective weights to the individual cues, and a mathematical model based on the weights

---

ed. Lewis R. Goldberg, "Simple Models or Simple Process? Some Research on Clinical Judgments," American Psychologist (July 1968), pp. 483-96.

could be created to make the predictions. A mathematical model could also be created by regressing an admissions officer's predictions against the cues or by using historical data. The decision could be made by using a mathematical model to provide an initial estimate of future success and then having the human decision maker adjust the estimates based on the subjective evaluation of the cues. All of the above methods of predicting future success, and there are others, involve a human decision maker, a mathematical model, or a human decision maker interrelating with the mathematical model. Both the human decision maker and the mathematical models bring advantages to the prediction process.

2.2.1 Advantages of the Mathematical Model

There are two major components of prediction accuracy: the selection of the information cues to be used in making the prediction (information choice), and the use of the selected cues (information use). Einhorn hypothesized that the major factor affecting prediction accuracy is the selection of the cues.[32] Einhorn's hypothesis has been supported by the empirical research of Abdel-khalik and El-Sheshai, who examined the effect of information choice on prediction accuracy and concluded that "the subjects' choice of infor-

---

[32] Hiller J. Einhorn, "A Synthesis, Accounting and Behavioral Science," Supplement to _Journal of Accounting Research_ (1976), pp. 167-79.

mation rather than their processing of the chosen cues was the limiting factor in predicting the environmental event."[33]

The inability of the human decision maker to select the optimal set of information cues for the prediction process creates a systematic error, reducing prediction accuracy. The source of the systematic error, information choice, can be eliminated by the use of a mathematical model for cue selection.

The second major component of prediction accuracy, cue usage, can be divided into its subcomponents by the lens model equation:[34]

$$r_a = R_e \, G \, R_s$$

The lens model equation provides insight into two ways to improve prediction accuracy. The first is the reduction of the random error caused by the inconsistent application of task knowledge (R_s). This random error can be reduced by replacing the decision makers with mathematical models based on their decisions, the decision maker's model.[35]

---

[33]The Abdel-khalik and El-Sheshai study is the only lens model research known to the author in which the subjects select their own cues. A. Rashad Abdel-khalik and Kamel M. El-Sheshai, "Information Choice and Utilization in an Experiment on Default Prediction," Journal of Accounting Research (Autumn 1980), p. 325.

[34]Kenneth R. Hammond and David A. Summers, "Cognitive Control," Psychological Review (January 1972), pp. 58-67.

[35]Ashton, Human Information Processing, pp. 34-43.

Second, prediction accuracy can be improved by reducing the systematic error which has been created by the decision maker's imperfect weighting of the cues (G). This systematic error can be eliminated from the prediction process by replacing the decision maker or the decision maker's model with the environmental model, which provides the optimal weighting of the cues.

2.2.2 Advantages of the Human Decision Maker

Although mathematical models may improve decision accuracy by eliminating random and systematic error from the decision making process, the human decision maker has two advantages over mathematical models. First, the mathematical model described above, and to be used in this research, is a linear model. Research, however, indicates that human decision makers use information in a curvilinear, configural, and sequential fashion.[36] Second, human decision makers bring a wealth of information that cannot be captured by a mathematical model into the decision making situation. While not mathematically quantifiable, their insight into the decision making environment may contribute significantly to decision accuracy.

2.3 LENS MODEL STUDIES

Discussed in this section are lens model studies, in

---

[36]Goldberg, "Simple Models or Simple Process?," p. 486.

both psychology and accounting, in which the primary emphasis is the measurement of prediction accuracy and the examination of the components of prediction accuracy. Lens model methodology has its roots in clinical psychology of the early 1960's. It was first employed by accounting researchers in the early 1970's. The results of some of the major studies of the cognitive psychologists will be discussed, and then a review of the literature in the area of accounting will be conducted.

2.3.1 Early Lens Model Research in Psychology

In 1968, Lewis R. Goldberg reviewed fifteen studies focusing on the accuracy of the clinical judgments of psychologists in predicting the outcome of the lives of their patients. He found that neither professional training and experience nor the amount of information affects the degree of judgment accuracy achieved by the expert judge. Goldberg states that:

> . . . the rather typical findings [are] that clinical judgments tend to be (a) rather unreliable . . . (b) only minimally related to the confidence and to the amount of experience of the judge, (c) relatively unaffected by the amount of information available to the judge, and (d) rather low in validity on an absolute basis . . .[37]

Goldberg summarizes his literature review:

> . . . over a large array of clinical judgment tasks (including by now some which were specific-

---

[37] Lewis R. Goldberg, "Simple Models or Simple Process?" p. 485.

> ally selected to show the clinician at his best and the actuary at his worst), rather simple actuarial formulae typically can be constructed to perform at a level of validity no lower than that of the clinical expert.[38]

An early study of the accuracy of clinical judgments was made by Goldberg in 1965.[39] Thirteen Ph.D.-level psychologists and sixteen advanced graduate students at the University of Minnesota evaluated the Minnesota Multiplasic Personality Inventory (MMPI) profiles of 861 psychiatric patients and diagnosed the patients as either psychotic or neurotic.[40] The MMPI profiles contained eleven scores, developed from a series of questions answered by the patient, that were used by the psychologists as cues in making their diagnoses. Each profile had been previously analyzed and approximately half of the patients classified as psychotic.

The judgment accuracy of the psychologists was low. The median judgment accuracy was $r_a = .28$. The most accurate judge achieved a judgment accuracy of $r_a = .39$, and the least accurate judge an accuracy of $r_a = .14$. The judgment accuracy of the average Ph.D. psychologist was

---

[38] Ibid.

[39] Lewis R. Goldberg, "Diagnosticians vs. Diagnostic Signs: The Diagnosis of Psychosis vs. Neurosis from MMPI," Psychological Monographs: General and Applied (1965), pp. 1-28.

[40] The data had been gathered by P. E. Meehl, Clinical Versus Statistical Prediction (Minneapolis: University of Minnesota Press, 1959).

equal to that of the average student. The accuracy of the environmental model was R_e = .46. Other correlations reported in the 1965 study were that of the composite judge, created by averaging the judgments of the twenty-nine judges, r_a = .35, and an environmental regression model generated using an unweighted composite of five MMPI scores, R_e = .44. The key result of this study is that the environmental models outperformed both the most accurate judge (r_a =.39) and the composite judge (r_a = .35).

In a 1970 study Goldberg continued his analysis of Meehl's data generating results that, at the time, were surprising; the judge's model outperformed the judge.[41] The model of each judge was created by regressing the judge's diagnosis against the eleven MMPI scores to generate a regression beta weight for each of the eleven scores. The beta weights constituted the judge's model. The accuracy of the judge's model, r_m, was measured by correlating the predictions of the model and the criterion diagnoses.

The median judgment accuracy of the judges' models was r_m = .33. The judgment accuracy of the most accurate model was r_m = .43 and of the least accurate model, r_m =.16. For twenty-eight of the twenty-nine judges, the accuracy of the judge's model was equal to or greater than that of the judge. When the judges' models were generated

[41] Lewis R. Goldberg, "Man Versus Model of the Man," pp. 422- 432.

using one-half the cases and the judgment accuracy generated using the remaining one-half, the judge's model outperformed the judge for twenty-five of the twenty-nine judges. Goldberg suggests the following conclusions:

> 1. Within the generality restriction obviously imposed by any one study, it has been demonstrated that linear regression models of clinical judges can be more accurate diagnostic predictors than are the humans who are modeled.
>
> 2. The composite judgement of all 29 clinicians, which was more accurate than that of the typical judge, was <u>not</u> improved by the modeling procedure.[42]

To summarize, Goldberg's 1965 and 1970 studies comparing the judgment accuracy of the environmental model, the judge, the model of the judge, and the composite judge resulted in a judgment accuracy of the environmental model that was the most accurate, $R_e = .46$. The judgment accuracy of the judge was the least accurate with a median accuracy of $r_a = .28$. The most accurate judge achieved a correlation of $r_a = .39$. The model of the judge was more accurate than the judge, but did not attain the judgment accuracy of the environmental model. The median accuracy of the model of the judge was $r_m = .33$, and the correlation of the most accurate model was $r_m = .43$. For twenty-five of the twenty- nine psychologists, eighty-six percent, the judge's model outperformed the judge, a statistically significant result with a p-value of $p = .001$. The judgment

---

[42]Ibid., p. 430.

accuracy of the composite judge was $r_a = .35$. There was no practical or statistically significant difference between the prediction accuracy of the median model of the judge and the composite judge.

Hammond and Summers provided a theoretical base for the results of Goldberg's research.[43] They proposed "that performance in cognitive tasks involves two distinct processes: the acquisition of knowledge and cognitive control over the knowledge already acquired."[44] When the variance in the criterion variable can be accounted for by a linear model, prediction accuracy (r_a) is a function of three factors; task uncertainty (R_e), task knowledge (G), and cognitive control (R_s). The relationships among the factors can be represented by the equation:

$$r_a = R_e \, G \, R_s$$

Task uncertainty (R_e) "sets the limit on the extent to which achievement (r^a) may occur, even if knowledge (G) and control (R_e) are perfect."[45] Knowledge (G) "measures the extent to which the subject correctly detected the properties of the task."[46] G denotes the degree to which the beta weights of the environmental model and the model

---

[43] Kenneth R. Hammond and David A. Summers, "Cognitive Control," pp. 58-67.

[44] Ibid.

[45] Ibid., p. 60.

[46] Ibid.

of the judge are isomorphic. When the judge has perfect knowledge, the beta weights of the environmental model and the judge's model will be the same and G will be unity (G = 1). Cognitive control (R_s) "measures the extent to which the subject controls the execution of his knowledge."[47] When R_s is less than unity, the judge is not applying knowledge in a consistent fashion and judgment accuracy is reduced.

In Goldberg's 1970 study, task median knowledge was G = .68 and cognitive control was R_s = .77.[48] According to Hammond and Summers, then, the environmental model outperformed the human judge because both task knowledge and cognitive control were less than unity. The model of the judge outperformed the judge because cognitive control was less than unity.[49]

The linear model has been found to be robust in predicting the judgments of the psychologists.[50] Dawes and

---

[47]Ibid.

[48]Goldberg, "Man Versus Model of Man", p. 426.

[49]Goldberg demonstrates mathematically that when the variance in the criterion variable can be accounted for by a linear model, prediction accuracy is a function of task uncertainty and cognitive control: $r_a = R_e G$. Ibid., p. 425.

[50]Slovic and Lichtenstein, in their in depth review of lens model studies, state, "This line of research, employing both correlational and ANOVA techniques, can be summarized simply and conclusively. . . . The linear model accounts for all but a small fraction of predictable variance in judgments across a remarkably diverse spectrum of tasks." Paul Slovic and Sarah Lichtenstein, "Comparison of Bayesian

Corrigan provide three reasons for the robustness of linear models.[51] First, linear models are good approximations of all multivariate models that have a conditionally monotonic relationship between the cues and the criterion event. "That is, the variables can be scaled in such a way that higher values on each predict higher values on the criterion, independently of the values of the remaining variables."[52] In Goldberg's research, for example, the higher the psychiatric patient scores on the schizophrenia scale or the paranoia scale, the higher the probability that the patient is psychotic, regardless of the scores of the other variables. Second, the relative beta weights of the cues are not affected by errors in the criterion variable. The correlation between the prediction of the environmental model (Y_e) and the criterion variable (Y_e) remains high even though there are errors in the measurement of the criterion. Third, error in the measurement of the cues tends to make the output of nonlinear models more linear. The net result is that the nature of the environment in which the lens model methodology is applied explains the high degree of accuracy of the linear models.

---

and Regression Approaches," p. 679.

[51]Robyn M. Dawes and Bernard Corrigan, "Linear Models in Decision Making," Psychological Bulletin (February 1974), pp. 95-106.

[52]Ibid., p. 98.

Dawes and Corrigan examined the relative accuracy of four different linear models using four data bases from previously published research. The first data set is from the 1965 Goldberg study presented above.[53] The second data set is from a study by Wiggins and Kohen in which eighty University of Illinois students were asked to predict the grade point averages of ninety first year students and forty-one graduate students from the University of Oregon.[54] The third data set is from a study by Dawes.[55] The faculty ratings of one hundred-eleven Ph. D. students in psychology at the University of Oregon were correlated with the ratings of the students by the admissions committee prior to their acceptance. The faculty ratings are the criterion variables and the ratings of the admissions committee are considered the predictions of the faculty ratings. The fourth data set is from a study by Yntema and Torgerson.[56]

---

[53]Goldberg, "Diagnosticians vs. Diagnostic Signs:," pp. 1-28.

[54]N. Wiggins and E. S. Kohen, "Man vs. Model of Man Revisited: The Forecasting of Graduate School Success," Journal of Personality and Social Psychology (July 1971), pp. 100-106.

[55]Robyn M. Dawes, "A Case Study of Graduate Admissions: Application of Three Principles of Human Decision Making," American Psychologist (February 1971), pp. 180-188.

[56]D. B. Yntema and W. S. Torgerson, "Man-Computer Cooperation in Decisions Requiring Common Sense," IRE Transactions of the Professional Group on Human Factors in Electronics (1961), pp. 20-26.

> Experimentors assigned values to ellipses presented to subjects on the basis of each figure's size, eccentricity, and grayness; . . . Subjects in this experiment were asked to estimate the value of each ellipse . . . The problem was to predict the true (i.e., experimenter assigned) value of each ellipse on the basis of its size, eccentricity, and grayness.[57]

The purpose of Dawe's study was to demonstrate that linear models achieve a higher judgment accuracy than the human judge. To achieve this purpose Dawes and Corrigan created, using the data from each of the four studies, in addition to the environmental model and the judge's model, a linear model in which the cues were randomly weighted, and models in which the cues were equally weighted. The signs of the cues used in the randomly weighted and the equally weighted models were selected a priori.

In four of the five data sets, the environmental model achieved the highest judgment accuracy followed by the equally weighted model. In the Yntema and Torgerson data set, the judgment accuracy of the environmental model and the equally weighted model were the same. In all cases the human judge was the least accurate judge. That the environmental model is the most accurate linear model is logical because it contains the optimal cue weights. The explanation of the superior accuracy of the judge's model over that of the judge provided by Dawes and Corrigan is that, in addition to the three reasons for the robustness

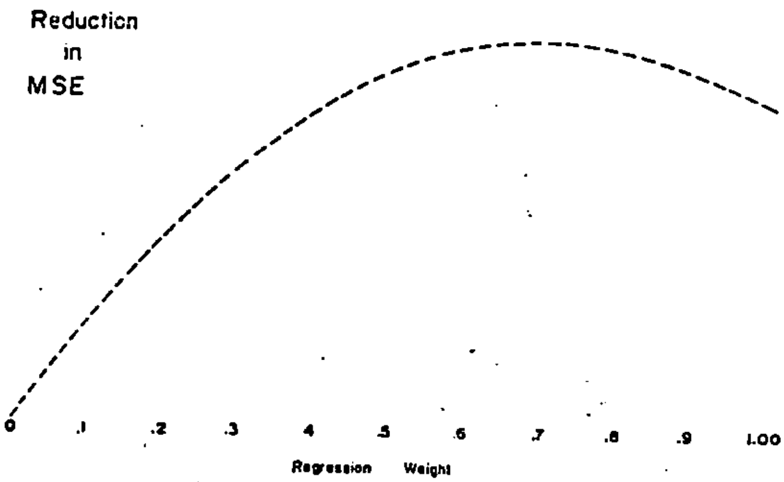
---

[57]Dawes and Corrigan, "Linear Models", p. 96.

of linear models described earlier, the linear models are robust over deviation from the optimal cue weights; the curve of the change in judgment accuracy resulting from a change in the beta weights of the cues is "flat". An example of the change in mean square error as a function of change of the beta weight of a single predictor linear model is replicated in Figure 3, p. 39. The optimal beta weight of the cue in Figure 3 is .71, and the curve at the upper range of the beta weights is nearly flat. As a result, because the expert judge "knows at least something about the direction of the variables, [the judge's model] yields weights near optimal."[58]

An interesting finding of the Dawes and Corrigan study was the relatively high degree of accuracy achieved by the equally weighted models. Dawes and Corrigan believe that the high degree of accuracy of the equally weighted models results because the number of observations of the criterion variables was too small to generate stable beta weights. They therefore believe that equally weighted models may be superior to the environmental models when the models are applied to new data. This conclusion is supported by the fact that in two of the four data sets in which the environmental model was cross validated using new data, the judgment accuracy of the equally weighted

[58]Ibid., p. 103.



Reduction in mean square error (MSE) as a function of the believed correlation coefficient.

Figure 3

CHANGE IN MEAN SQUARE ERROR

Robyn M. Dawes and Bernard Corrigan, "Linear Models in Decision Making," Psychological Bulletin (February 1974), p. 104.

model was superior.

Dawes and Corrigan conclude:

> Linear models work because the situations in which they have been investigated are those in which: (a) The predictor variables have conditionally monotone relationships to criteria (or may easily be rescaled to have such a relationship); (b) there is error in the dependent variable; (c) there is error in the independent variables; and (d) deviations from optimal weighting do not make much practical difference. These situations abound.[59]

To summarize, Goldberg, in his two studies, found the judgment accuracy of the human judge to be quite low. Judgment accuracy was improved by replacing the human judges with their models. The highest judgment accuracy was achieved by the environmental model. It was shown mathematically that the major reasons for the superiority of the environmental model were a lack of knowledge of the task and a lack of cognitive control over the knowledge acquired. Hammond and Summers provide the theoretical explanation for the relationship between task knowledge and cognitive control and judgment accuracy. Dawes and Corrigan provide additional theoretical explanation for the empirical support of the robustness of the linear models. The results of these studies have been widely supported by other research studies. Libby, in a comprehensive review of the clinical versus statistical prediction literature of the cognitive psychologists summarizes his observations:

---

[59]Ibid., p. 105.

> In many important decision-making situations, the environmental predictability of the available information is low. However, even in the situations where environmental predictability is relatively high, poor judgmental achievement is the norm.
>
> Both human inconsistency and misweighting of the cues contribute to poor achievement. Combining quantitative information in repetitive tasks does not appear to be a function that people perform well. Thus, in these situations, replacing people with statistical models (e.g. environmental regression models, models of the man, and equal weighting models) shows promise for increasing predictive accuracy.[60]

2.3.2 Accounting Studies Relating to Prediction Accuracy

The major lens model studies in accounting relate to one of two areas. They examine either bankruptcy prediction or the prediction of stock prices.

2.3.2.1 Prediction of Bankruptcy

Libby, in a 1975 study, had forty-three commercial loan officers evaluate sixty cases, consisting of five financial ratios and representing actual firms, half of which had failed within the previous three years.[61] The task of the participants was to examine each case and classify the firm as failed or nonfailed and to record confidence in the prediction on a three point scale. The participants were informed that half of the firms had failed. Prediction accuracies were measured by scoring +1 for each

---

[60]Robert Libby, Accounting and Human Information Processing, pp. 28-29.

[61]Robert Libby, "Accounting Ratios and the Prediction of Failure: Some Behavioral Evidence," Journal of Accounting Research (Spring 1975), pp. 150 -161.

correct prediction and -1 for each incorrect prediction.[62]

Using a binomial test, it was determined that the probability of randomly predicting thirty-seven or more cases correctly was slightly less than .05. The prediction accuracy of the environmental model was fifty-one; fifty-one of sixty cases were correctly classified as bankrupt/non-bankrupt by the environmental model. The mean prediction accuracy of the participants was 44.4. The most accurate participant achieved a prediction accuracy of fifty and the least accurate participant had a prediction accuracy of twenty-seven. Three participants had a prediction accuracy below thirty-seven. A composite judge was created by computing the response for each firm, fail or not-fail, across all participants. If the average response was less than zero, the firm was considered failed; and if the average response was greater than zero, the firm was considered nonfailed. The composite judge correctly classified forty-nine cases, only one less than the best judge and two less than the environmental model. To summarize, the environmental model achieved the highest prediction accuracy, fifty-one, followed by the composite judge, forty-nine. The human judge was the least accurate predictor with an achievement of 44.4.

---

[62]Hit rates are used in this study instead of the correlation coefficients used by Goldberg. Goldberg, "Simple Models," pp. 483-496.

In addition to prediction accuracy, Libby measured judgment consensus and judgment stability over time and confidence. Consensus is the agreement of all pairs of participants as to a firm's failed/nonfailed status. The consensus ranged from thirty-one to fifty-seven and averaged forty-eight. Judgment stability over time was determined by splitting the sixty cases into two individual case sets, the first case set containing thirty firms and the second containing forty firms, ten of which were repeats. The two sets of cases were then completed in two separate sittings one week apart. Thirteen of the participants exhibited perfect stability; ten of the ten repeat cases were consistently rated failed/nonfailed. The mean stability was 8.9. Libby found no significant correlation between the participant's reported level of confidence and actual performance. In 1976, Libby reported a further analysis of the data generated in the 1975 research discussed above.[63] Libby generated the participants' models and compared the accuracy of the models with that of the participants. The participants achieved a mean prediction accuracy of 44.4, and the mean prediction accuracy of the models was 43.3. Thirty-three of the forty-three participants performed as well as or better than their models. In only

---

[63]Robert Libby, "Man Versus Model of Man: Some Conflicting Evidence," Organizational Behavior and Human Performance (June 1976), pp. 23-26.

ten instances did the model outperform the participant. These results are statistically significant and are at variance with other studies of bootstrapping research.[64] Libby found a significant negative correlation between the prediction accuracy of the judge and the incremental prediction accuracy of the model over the judge; as the prediction accuracy of the judge increases, the increase in prediction accuracy attained by replacing the judge with the model of the judge decreases. "Of the 21 judges whose validity was 45 to 60 or greater, 20 were at least as valid as their models. This would suggest that bootstrapping does not improve the performance of the most accurate judges."[65]

Three differences between Libby's study and previous research on bootstrapping were presented as possible reasons for the results. First, the criterion variable, business failure, was well defined and reliably measured, and the criterion variables of prior research were not well defined or reliably measured. Second, the commercial bankers were experts at the task, much of their working time is spent performing financial analysis. The participants in the prior studies were not specialists at the experimental

---

[64]Bootstrapping occurs when the prediction accuracy of the decision maker's model is greater than that of the decision maker.

[65]Libby, "Man Versus Model of Man," p. 9.

task. Third, the abnormal nature of some of the cases created skewed cue distribution. The relationships between the cues and the criterion variable were not monotonic.[66] Libby concluded that the participants were correctly using nonlinear models.

Zimmer replicated Libby's study using thirty loan officers from two Australian banks. Five financial ratios to predict failure/nonfailure of forty-two disguised industrial firms, half of which had failed one year subsequent to the date of the financial ratios, were used.[67] The participants were told that "approximately half" of the firms had failed. The results were consistent with those of Libby's study. The environmental model correctly predicted failure or nonfailure in thirty-seven of forty-two firms. The mean prediction accuracy of the participants

---

[66]Goldberg rescaled Libby's skewed cues by a normalizing transformation and repeated all analyses with a dramatic change in the results. The model outperformed the judge for 72% of the participants versus 27% in the original study. Goldberg concludes that Libby's results do not conflict with those of previous studies. The models of the judges performed poorly because the skewed cues created erroneous regression weights and the presence of extreme values of the cues caused a skewed distribution of the predicted criterion variables. Lewis R. Goldberg, "Man Versus Model of the Man," pp. 13-22. Libby rebuts Goldberg stating that his analysis is flawed and that he misinterprets Libby's original findings. Robert Libby, "Man Versus Model of Man: Need for a Nonlinear Model," Organizational Behavior and Human Performance (June 1976), pp. 23-26.

[67]Ian Zimmer, "A Lens Study of the Prediction of Corporate Failure by Bank Loan Officers," Journal of Accounting Research (Autumn 1980), pp. 629-636.

was 32.4 with twenty-eight of the participants achieving greater than random prediction accuracy. Again, the binomial theorem generated a ninety-five percent confidence level of a prediction rate greater than random if the participant correctly predicted the status of twenty-seven firms. The composite judge made thirty-six correct predictions. Zimmer did not report the model of the man. These results are in conformance with those of Libby's study: the highest prediction accuracy was achieved by the environmental model, followed by the composite judge. The human decision maker was the least accurate predictor.

Zimmer also examined the participant's degree of confidence and determined that "accuracy improved with greater confidence."[68] In regard to judgment consensus, the mean consensus between the loan officers was 30.35 indicating that the loan officers' predictions were reasonably homogeneous. Finally, the research was replicated using upper level accounting students from the University of New South Wales:

> Results indicated that the responses of students and loan officers were not significantly different. Students performed at better than random accuracy, . . . showed a high degree of judgmental consensus, and the majority were outperformed by the consensus judge. . . . Clearly, inferences drawn from this experiment using loan officers would have been identical if students had been

---

[68]Ibid., p. 633.

used as surrogates.[69]

Like Libby, Zimmer reported the comparison of man versus model of the man in a separate study.[70] Discriminant models were developed from the forty-two predictions for each of the thirty participants. A significant statistical difference was found between the prediction accuracy of the loan officers and their models; the mean prediction accuracy of the loan officers was 32.4, compared to 33.7 for the models. The models were cross-validated using the Lachenbruch hold-out technique which has the effect of comparing the loan officers' predictions to that of their models created from the cases held out. The mean prediction accuracy of the cross-validated models was 33.1 and was not statistically different from the mean prediction accuracy of the participants.

Casey performed a partial replication of the 1975 Libby study using forty-six bank loan officers to predict which of thirty firms would become bankrupt within three years.[71] Casey made two major departures from Libby's experimental design. First, the frequency of failure was

---

[69]Ibid., p. 635.

[70]Ian Zimmer, "A Comparison of the Prediction of Loan Officers and Their Linear-Additive Models," Organizational Behavior and Human Performance (February 1981), pp. 69-73.

[71]Cornelius J. Casey, Jr., "The Usefulness of Accounting Ratios for Subjects' Predictions of Corporate Failure: Replication and Extensions," Journal of Accounting Research (Autumn 1980), pp. 603-613.

not disclosed to the participants. (Half of the firms were bankrupt.) The second major departure was that three consecutive years of financial ratios were presented for each of the thirty firms prior to the prediction date, thus providing the participant with an opportunity to observe a firm's deteriorating condition prior to failure. Again, the task was to predict which firms would become bankrupt within the subsequent three year period.

The average prediction accuracy of the judges was low; seventeen of thirty firms were correctly classified. Thirteen of the nonbankrupt firms were correctly classified; however, only four of the fifteen bankrupt firms were correctly cassified. "Whereas 41 of the 46 subjects could predict the nonbankrupt firms with an accuracy greater than 50 percent, no subject could do the same for the bankrupt firms."[72] The composite judge, computed by totaling the predictions of the forty-six participants, correctly classified eighteen of the thirty firms. The composite judge correctly classified fourteen of the fifteen nonbankrupt firms and only four of the bankrupt firms. There was no statistical difference between the prediction accuracy of the composite judge and the participants. The results suggest to Casey that because the participants were not informed of the distribution of failed/nonfailed firms

[72]Ibid., p. 609.

in the sample, they injected their own low prior probabilities of failure causing the number of firms being identified as bankrupt to be small. Only twenty percent of the participants' and seventeen percent of the composite judge's predictions were for failure.

Casey also examined the relationship between the participants' confidence level and their achieved prediction accuracy. Confidence was defined as the participant's estimate of the number of correct predictions. As was the case in Libby's study, there was no significant relationship between confidence level and overall prediction accuracy.

A study by Kida[73] was motivated by Altman and McGough's finding that a discriminant bankruptcy model was able to predict the failure of eight-two percent of a sample of firms, while auditors mentioned going-concern problems in the audit report for only forty-four percent of these firms.[74] Kida argued that the consequences of qualifying a firm's financial statements, losing the client, or not qualifying the statement, lawsuits by investors and creditors, as well as the perceived likelihood of going-concern

---

[73]Thomas Kida, "An Investigation into Auditors' Continuity and Related Qualification Judgments," Journal of Accounting Research (Autumn 1980), pp. 506-523.

[74]E. Altman and T. McGough, "Evaluation of a Company as a Going Concern," Journal of Accountancy (December 1974), pp. 50-57.

problems, impacted on the auditor's decision to issue a qualified opinion. The purpose of the research was, therefore, to first determine the auditor's accuracy in predicting going-concern problems compared to that of the environmental model, and second, identify the other factors in the decision to qualify or not qualify the opinion, if the auditor's prediction accuracy is greater than that of the qualified opinions given.

In the first phase of the study, twenty-seven audit partners from a national CPA firm used five financial ratios to predict which of forty firms had going-concern problems. A firm was determined to have going-concern problems if:

> it was likely that the firm would experience one or more of the following events within one year: (a) enter receivership; (b) enter reorganization proceedings; (c) inability to meet interest payments; (d) experienced its third consecutive year of substantial losses; (e) liquidate its assets; (f) experience its third consecutive year of significant deficits.[75]

Half of the firms were determined to have going-concern problems. The participants were informed that some of the firms fit the going-concern definition but "there were not necessarily an equal number in each category."[76]

A discriminant model correctly classified thirty-six of the forty firms. The average number of correct responses by the auditors was 33.2. The range was from twenty-four

---

[75] Kida, "Qualification Judgments," p. 509.

[76] Ibid., p. 511.

to thirty-seven with twenty-five of the twenty-seven auditors achieving a prediction accuracy greater than chance accuracy. It was determined, using a binomial distribution, that a prediction accuracy of twenty-eight or greater had a less than one percent chance. The mean prediction accuracy of the linear models of the auditors was 33.7, with a range of thirty to thirty-nine correct responses. Sixteen of the auditors had a prediction accuracy equal to or greater than their models. There was no statistical difference between the accuracy of the auditors and that of their models. To summarize, the prediction accuracy of the environmental model was ninety percent, the mean prediction accuracy of the models of the auditors was eighty-four percent, and the mean prediction accuracy of the auditors was eighty-three percent. The auditors were able to predict going-concern problems with a high degree of accuracy.

The purpose of the second phase of the study was to elicit the auditors' qualification decisions and to determine the attitudinal factors that affected the decisions. The auditors indicated that they would issue a qualified opinion or a disclaimer for an average of 13.2 firms while recognizing going-concern problems for an average of 17.5 of the forty firms. The auditors qualified or disclaimed seventysix percent of the firms determined to have going-concern problems. Kida was able to link the propensity to qualify to specific perceptions about the consequences of qualifying; loss of the client, or not qualifying, lawsuit by the client's creditors.

In addition to the above findings, Kida found no statistically significant relationship between the auditors' confidence in their going-concern judgments and their prediction accuracy. A high degree of consensus was found among the auditors. "It would appear, . . . that a substantial level of agreement exists between auditors when they make continuity decisions from financial statement data."[77]

Abdel-Khalik and El-Sheshai performed a lens model bankruptcy study in which the participants selected their own cues.[78] The purpose of the study was to evaluate the effects of information choice, as well as the effects of information processing, on prediction accuracy. Three methods of processing information were examined; the environmental model, the model of the man, and the human participant. Each of these methods of prediction can use cues chosen either by a mathematical model or by the human decision maker. The result is six combinations of information selection and processing which are shown in Table 1, p. 53.

Twenty-eight commercial loan officers were asked to

---

[77]Ibid., p. 519.

[78]A. Rashad Abdel-Khalik, "Information Choice and Utilization in an Experiment on Default Prediction," Journal of Accounting Research (Autumn 1980), pp. 325-341.

Table 1

STRATEGIES OF PREDICTIONS

[illegible]iction of Environmental Event (Default)	Information Selection	
	by Human	by Mathematical Models
[illegible] human	(1) $HP \mid HS$	(4) $HP \mid MS$
By "models of man"	(2) $MP_s \mid HS$	(5) $MP_s \mid MS$
By environmental models (optimal)	(3) $MP_e \mid HS$	(6) $MP_e \mid MS$

* H = human; M = model; S = selection [illegible] choice; P = prediction; the subscripts refer to the basis [illegible] generating the model as to predicting human response (s) or the environmental event (e).

$HP \mid HS$ is [illegible] prediction of the environmental event by human subjects based on [illegible] information sel[illegible]. This is frequently called the achievement index.

$MP_s \mid HS$ is the prediction of the environmental event using the mechanical model built to struc[illegible] [illegible]n responses in a linear form using hu[illegible] [illegible] information.

[illegible] MS is the model's environmenta[illegible] using the model-selected information

A. Rashad Abdel-Khalik and Kamal M. El-Sheshai, "Information Choice and Utilization in an Experiment on Default Prediction," <u>Journal of Accounting Research</u> (Autumn 1980) p. 329.

predict which of thirty-two firms would default on loans. (Half of the firms had defaulted.) The participants were informed only that "the proportion of defaulted to nondefaulted firms was not representative of the true proportion observed in real life."[79] Default on loans was selected as a criterion variable because it was of more direct concern to loan officers (more firms default on debt than go bankrupt) and because default normally precedes bankruptcy by more than a year. The human selection of the cues was effected by asking the loan officers to select a maximum of eight cues from ten financial ratios that were used in previous research and eight trends, the change in a financial item during the previous five years. The mathematical selection of the cues was effected using step-wise discriminant analysis.

The results, based on the percentage of correct predictions, are presented in Table 2, p. 55. Several aspects of the results are worth close examination. First, the models of the participants (the model of the man) and the environmental mathematical model using information selected by the humans, did not perform significantly better than the participants. No significant increase in prediction accuracy was achieved by removing the participant from the processing of information. There was, however, a

[79]Ibid., p. 331.

Table 2

RESULTS OF PREDICTION STRATEGIES

Strategy #	Denoted	Evaluating the Predictive Ability of the Environmental Event	% of Correct Prediction (Hit Rate)
S1	$HP \mid HS$	The judge's validity (achievement index) using human-selected information.	0.625 (average)
S2*	$MP_h \mid HS$	The validity of the "model of man" using human-selected information.	0.625 (average)
S3	$MP_e \mid HS$	Environmental mathematical model using information selected by humans.	0.675
S4	not studied		
S5	not studied		
S6	$MP_e \mid MS$	Environmental mathematical model using model-selected information.	0.906

* Strategy S2 uses the linear model estimated from human responses and the information cues that they used to predict the environmental event. The linear model estimated from human response (the model of man) correctly predicted 0.84 of the responses of the judges. The 0.625 correct classification shown in the table is the percentage of correct prediction of the environmental event using the "model of man."

[9] One main reason for breaking the sample was our concern that having subjects process too many information cues in the first round would be incongruent with the evidence on limited cognitive ability (Tversky and Kahneman [1974]).

A. Rashad Abdel-Khalik and Kamal M. El-sheshai, "Information Choice and Utilization in an Experiment on Default Prediction," Journal of Accounting Research (Autumn 1980), p. 336.

significant increase in prediction accuracy as a result of using a mathematical model to select the cues.

> . . . it appears that human predictions of default fell short of the predictive ability of the mathematical models used here because of the less than optimal use of information cues. Neither the processing nor the weighting of the cues appeared to contribute significantly to this lower performance.[80]

In addition to comparing the mean prediction accuracy of the participants to that of the environmental model using model-selected information, Abdel-Khalik and El-Sheshai compared the participants' prediction accuracy to five other benchmarks:

1. A multivariate discriminant model containing the four variables most frequently selected by the loan officers resulted in a seventy-five percent accuracy in the classification of the firms.
2. The current ratio correctly classified eighty-four percent of the firms. The rule of thumb for applying the current ratio is to classify a firm as defaulted if the current ratio is equal to or less than two.
3. The debt/total assets ratio correctly classified seventy-two percent of the firms. The rule of thumb for applying the debt/total assets ratio is to classify a firm as in default if the ratio

---

[80]Ibid., p. 337.

is equal to or less than one.

4. A multivariate discriminant model containing the three ratios that Beaver found to be the most predictive of failure (cash flow/total debt, net income/total assets, and total debt/total assets)[81] correctly classified the firms with seventy-two percent accuracy.
5. A multivariate discriminant model generated from a validation sample correctly classified seventy-eight percent of the firms.

All five of the benchmarks achieved prediction accuracies greater than the average prediction accuracy of the participants. A ninety percent confidence level generates an upper bound of 0.76. The prediction accuracy current ratio and the environmental model generated from the validation sample fell above the upper bound of the confidence interval and are, therefore, considered statistically significant. It is important to note that the environmental models using Beaver's three most predictive ratios and the ratios from the validation sample are the result of samples that are independent of the thirty-two firms used in the study. To summarize, mathematical models, even simple ones containing one ratio, consistently outper-

---

[81]Willian H. Beaver, "Financial Ratios as Predictors of Business Failure," Supplement to <u>Journal of Accounting Research</u> (1976), pp. 196-206.

formed the human decision makers.

2.3.2.2 Stock Price Prediction Studies

In 1977, Wright had thirty-nine second-year MBA students from the University of California at Berkeley predict both the price change and the percentage change in price of the common stocks of sixty actual firms.[82] To make the predictions, the participants were given seven cues drawn from the COMPUSTAT data file, containing both accounting and capital market information. The primary purpose of the research was to measure the prediction accuracy of the participants.

The accuracy of the environmental regression model in predicting the change in stock prices was $R_e = .58$. The median prediction accuracy of the participants' predictions of the change in stock prices was $r_a = .16$, with a maximum prediction accuracy of $r_a = .44$ and minimum of $r_a = -.08$. The prediction accuracy of the environmental model for predicting the percentage change in stock prices was $R_e = .49$. The median prediction accuracy of the participants' predictions of the percentage change in stock prices was $r_a = .20$, with a maximum prediction accuracy of $r_a = .46$ and a minimum of $r_a = -.12$. The overall prediction accuracy of the participants was quite low with only

[82] Willian F. Wright, "Financial Information Processing Models: An Empirical Study," The Accounting Review (July 1977), pp. 676-689.

forty percent of the participants achieving a prediction accuracy greater than chance (a = .05).

In addition to measuring prediction accuracy, Wright examined cognitive biases, judgment consensus, the participants' use of cognitive models, and the use of nonlinear decision models. Wright tested for a cognitive bias on low versus high priced securities and found that the participants "consistently overestimated the prices of low-priced securities, while there was no indication of bias for relatively high-priced stocks."[83] Interjudge consensus was determined by the product-moment correlations between a participant's judgments and those of the other participants. Little interjudge consensus was found. This finding is consistent with the substantial heterogeneity observed in the predictions of the individual participants. The range of the estimated mean stock price change estimated by the participants was $1.96, with a low mean estimate of -$4.63 and a high of $9.99. Cluster analysis revealed four distinct types of congitive models were used by the participants. Wright, however, did not delineate the heuristics used. Finally, it was determined that as a group the participants "used congitive models well-approximated by the linear model."[84]

---

[83]Ibid., p. 682.

[84]Ibid., p. 683.

To summarize, the overall prediction accuracy of the participants was low with only approximately forty percent of the participants able to predict with better than random accuracy. The dispersion of predictions resulted in low interjudge consensus and provided evidence of numerous information processing models. Cluster analysis confirmed the use of numerous models, with four distinct cognitive models being revealed. Finally, as a group, the participants' decisions indicated the use of linear models.

Wright, in a continuation of the above study, examined the participants' self-insight.[85] The participants' subjective weighting of the cues was captured by having them allocate one hundred points among the seven cues according to the relative weights they gave to each cue in making their predictions. Self-insight was tested in two ways. First, the degree of correlation between the subjective cue weights and the weights derived from the multiple linear regression equations representing the participants' models was determined. Second, the subjective cue weights were used in a linear model to predict the participants' judgments, which correlated with the actual judgments. If the participants understand the relative importance of the individual cue weights in the generation of their pre-

---

[85] Willian F. Wright, "Self-insight Into the Cognitive Processing of Financial Information," <u>Accounting Organizations and Society</u> (1977), pp. 323-331.

dictions, the participants have self-insight; the predictions of the participants' models and the subjective models should be highly correlated.

The mean (median) product-moment correlations of the subjective cue weights and the beta weights from the participants' models were .415 (.345). "These association statistics imply generally significant subject awareness of specific patterns of relative cue emphasis for this group of individuals."[86] The mean squared correlation of the outputs of the participants' models and of the subjective models was .28. The squared correlation was used because it indicates the percentage of the participants' predictions that are explained by the linear model. An equally weighted model generated a squared correlation of .17. The participants' own models generated a squared correlation of .43. Wright summarizes the results. "When the results reported here are compared with previous studies in other areas, an encouraging degree of self-insight into relative weighting of financial information is indicated."[87]

Wright replicated his 1977 study by having forty-seven first and second year graduate students use four cues to predict the percentage price change for fifty common stock

---

[86]Ibid., p. 327.

[87]Ibid., p. 329.

securities during the subsequent year.[88] The dimensions of cognitive processing examined included prediction accuracy, linear versus nonlinear processing of the cues, interjudge agreement, self-insight into the weighting of the cues, and the accuracy of the composite judge.

In conformance with Wright's previous study, the prediction accuracy of the participants was low. The median prediction accuracy of the thirty-five first year graduate students (Group One) was r_a = .20, and the prediction accuracy of the twelve second year graduate students (Group Two) was slightly better, r_a = .31. A prediction accuracy of r_a = .24 was considered significant at the .05 level of significance. Thirty-seven percent of the Group One participants and seventy-five percent of the Group Two participants achieved a prediction accuracy greater than chance. The median environmental predictability was R_e = .53. Wright disaggregated prediction accuracy into its systematic and random error components. The median matching index for the Group One (Group Two) participants was G = .48 (G = .46) with an interquartile range of [.29,.70] ([.43, .65]). Wright concludes that the participants "recognized and utilized environmental relationships."[89]

[88] Willian F. Wright, "Properties of Judgment Models in a Financial Setting," Organizational Behavior and Human Performance (February 1979), pp. 73-85.

[89] Ibid., p. 78.

The median consistency index for Group One (Group Two) was $R_s = .53$ ($R_s = .62$) with approximately half of the coefficients between .38 and .57 (.51 and .65). The Group Two participants employed their cognitive models in a more consistent manner than those of Group One. The final component of prediction accuracy examined was the nonlinear usage of the cues. The nonlinear cue utilization for Group One (Group Two) was $C = .03$ ($C = .20$); ". . . there is little evidence here for appropriate systematic nonlinear cue usage, and this conclusion is consistent with previous results."[90]

The prediction accuracy of the composite judge for Group One (Group Two) was .32 (.38), a significant increase in prediction accuracy over the median individual performance. The systematic and random error was diversified out of the composite model. Wright found moderate interjudge agreement and moderate self-insight into the cue weighting for the Group One participants and a higher level of self-insight for the Group Two participants.

To summarize the key points of the study, a low level of prediction accuracy was achieved by the participants as a result of low environmental predictability and a high degree of random and systematic error in the participants' processing of the cues. The composite judge significantly

---

[90]Ibid.

outperformed the median prediction accuracy of the participants, but not the prediction accuracy of the environmental model.

In a 1978 study, Ebert and Kruse examined the feasibility of applying bootstrapping to predicting future returns on common stock securities.[91] Five security analysts used twenty-two cues to predict the returns of thirty-five stocks. Five sets of bootstrapping models were created for each of the participants by randomly selecting twenty securities and, because of the large number of cues involved, using stepwise regression to estimate the participants' models. The models were then cross-validated on the fifteen remaining securities generating the participant's prediction accuracy, the prediction accuracy of the participant's model, the prediction accuracy of the environmental model, and the prediction accuracy of the composite judge for each participant for each of the five models. Two additional models, a stepwise regression model and a regression model using all twenty-two cues, were generated for each participant incorporating all thirty-five securities.

In eighteen of twenty-five instances using the cross-validated models, the participant's model achieved a higher prediction accuracy than the participant. In nineteen

---

[91] Ronald J. Ebert and Thomas E. Kruse, "Bootstrapping the Security Analyst," Journal of Applied Psychology (February 1978), pp. 110-119.

instances the environmental models outperformed the participants. There was no consistent pattern of superiority of the environmental models over the participants' models. The composite judge was generated for all seven models, and the prediction accuracy of the composite judge was higher than that of the participants in thirty-four of the thirty- five instances. The performance of the composite judge was superior to that of the environmental models in every instance. In general, the results are in conformance with those of the research studies previously discussed; the participants used linear models, and the participants were outperformed by the environmental model, their own models and the composite judge.[92]

2.4 INFORMATION OVERLOAD STUDIES

In 1956, Miller reviewed a large body of psychological research relating to the techniques for increasing the accuracy of human judgment.[93] He stated:

> The three most important of these devices are (a) to make relative rather than absolute judgments . . . (b) to increase the number of dimensions along which the stimuli can differ; or (c) to arrange the task in such a way that we make a sequence of several absolute judgments

---

[92]The authors state that, "the analysts' models imply a highly consistent, linear judgmental strategy." Ibid., p. 114.

[93]George A. Miller, "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," The Psychological Review (March 1956), pp. 81-97.

in a row.[94]

It is the second technique, to increase the number of dimensions along which stimuli can differ, that is of concern here.

Two terms need to be defined. The first, bit, is the amount of information required to make a judgment about two equally likely alternatives. One bit of information is required to determine if a coin is heads up or tails up. Two bits of information enable a judgment about four equally likely alternatives, four bits a judgment about eight equally likely alternatives, as so on. The second, channel capacity, is the greatest amount of information observers can receive, differentiate, and accurately integrate into their judgments. Channel capacity is measured in bits.

In his literature review, Miller found that the range of channel capacity for one-dimensional judgments (tones, loudness, visual position between two markers) is between 1.6 and 3.9 bits which equates to a range of four to ten categories. In a study by Pollack, for example, the upper limit of channel capacity was 2.5 bits when the participants were asked to identify tones.[95] The listeners could

---

[94]Ibid., p. 90.

[95]I. Pollack, "The Assimilation of Sequentially Encoded Auditory Displays," American Journal of Psychology (1953), pp. 745-749.

identify only six different pitches of sound without becoming confused.

The channel capacity increases with multidimensional stimuli. Miller reports a study by Pollack and Ficks.[96]

> . . . Pollack and Ficks conducted a test on a set of tonal stimuli that varied in eight dimensions, but required only a binary decision on each dimension. With these tones they measured the transmitted information at 6.9 bits, or about 120 recognizable kinds of sounds.[97]

While the addition of independent variable attributes (frequency, intensity, and length duration for acoustic variables) increases the channel capacity, the increase is at a decreasing rate, leading Miller to speculate that channel capacity will be maximized with seven to ten variable attributes.[98] To summarize, a human observer appears to be able to distinguish approximately seven categories within a unidimensional stimuli. The number of categories the observer is able to distinguish increases as new dimensions of the stimuli are added to an apparent maximum of seven to ten dimensions.[99]

---

[96]I. Pollack and L. Ficks, "Information of Elementary Multidimensional Auditory Displays," Journal of Accoustics Society of America (1954). 155-158.

[97]Miller, "The Magical Number Seven," p. 89.

[98]Ibid., p. 96.

[99]To relate Miller's literature review to the current study, the human decision maker can differentiate approximately seven categories within a stimuli (cue) and the number of stimuli that can be integrated into the judgment process is limited (seven to ten cues).

Dudycha and Naylor, in a 1966 study, examined the impact of increasing the number of environmental cues from one to two on prediction accuracy of human decision makers.[100] Six experimental conditions were created by adding to each of two first cues, with validities of $r_{e1} = .40$ and $r_{e1} = .80$, second cues with validities of $r_{e2} = .20$, $r_{e2} = .40$, and $r_{e2} = .60$. Each of the second cues were orthoginal to the first. The six experimental conditions created environmental models with the following correlations (R_e values); .998, .899, .819, .728, .549, and .457. Two hundred trials were created for each condition using computer simulation to generate the cues and the criterion variables. Both the cues and criterion variables generated consisted of two-digit numbers.

Sixty elementary psychology students were randomly assigned to a condition and completed the two hundred trials for each condition. "A single trial consisted of (1) displaying the two cues, (2) allowing time for all subjects to record their response, and (3) the display of the criterion ('correct' response) with the two cues visible."[101]

Learning appears to have taken place during the first one hundred trials after which prediction accuracy stabi-

---

[100] Linda W. Dudycha and James C. Naylor, "Characteristics of the Human Inference Process in Complex Choice Behavior Situations," Organizational Behavior and Human Performance (September 1966), pp. 110-128.

[101] Ibid., p. 118.

lizes. Generally, when a cue with low validity is paired with a cue with high validity the participant's prediction accuracy increases, and when a high validity cue is paired with a low validity second cue prediction accuracy decreases.

> From the base line achievement level of Condition .40, r_a is virtually unaffected with the addition of a cue with r_{e2} = .20, r_{a2} increased somewhat with the addition of a cue with r_{e2} = .40, and increased most with a second cue added whose validity equalled .60. However, from the base line achievement level of Condition .80 a different picture emerges. With the addition of a second cue with validity .20, performance is severely decreased from the base line. With r_{e2} = .40 the decrement is not so severe. With r_{e2} =.60 r_a does exceed the base line level, but only by a small magnitude.[102]

Dudycha and Naylor also found that the consistency with which the participants used their own models (R_s) decreased when a second cue with a low validity was added to a first cue with a high validity. The consistency index decreased substantially when a second cue with a validity of r_{e2} = .20 was added to a first cue with a validity of r_{e1} = .80. A less substantial decrease in the consistency index was observed when a second cue with a validity of r_{e2} = .40 was added, and the consistency index increased slightly when a second cue with a validity of r_{e2} =.60 was added. Consistency increased when a second cue was added to a low validity first cue. When the initial cue

---

[102] Ibid., p. 122.

was r_{e1} = .40, consistency increased slightly with the addition of a second cue of $r_{e2} = .20$ and $r_{e2} = .40$. There was a marked increase when cue $r_{e2} = .60$ was combined with $r_{e1} = .40$. The effects of cue validities on the matching index (G) were also examined. "The amount of matching was less when the second-cue validity equaled .20 than when it equaled .40 and .60; and less when the first cue validity equaled .80 than when it was .40."[103]

To summarize, the participants' prediction accuracy increased when a second high validity cue was added to a first low validity cue. Prediction accuracy decreased when a low validity cue was added to a high validity cue. The same pattern was observed with the consistency index. There was no distinct pattern to the matching index.

In a 1971 study, Einhorn examined the effects of varied amounts of information on prediction accuracy.[104] Faculty members were randomly assigned to three groups receiving either two, four, or six cues. The task was to rank two sets of fifteen hypothetical candidates to a graduate school of psychology. The Spearman rank order coefficient was used to measure the correlation between the participants' predictions and the actual rankings for both data sets.

---

[103] Ibid., p. 125

[104] Hellel J. Einhorn, "Use of Nonlinear, Noncompensatory Models as a Function of Task and Amount of Information," Organizational Behavior and Human Performance (January 1971) pp. 127.

Information overload was found in both data sets after the first two cues. The participants' prediction accuracies for the first set of hypothetical applicants was $r_a = .885$ for two cues, $r_a = .748$ for four cues, and ra = .750 for six cues. The participants' prediction accuracies for the second set of hypothetical applicants was $r_a = .920$ for two cues, $r_a = .817$ for four cues, and $r_a = .590$ for six cues.

The first accounting study concerning information overload to be reviewed is a 1972 study by Barefield which examines the impact of disaggregated versus aggregated data on the prediction accuracy of human decision makers.[105] To avoid confounding the effects of the degree of information content and the degree of data aggregation, accounting reports were generated that had the same information content. The reports were considered to have the same information content when the environmental predictability of the reports was the same.

The experimental task was to examine labor efficiency and/or material quantity variances and to determine if labor was in or out of control. The probability of the labor being out of control was fifty percent and the participants were informed of the prior probability. Twenty-

---

[105]Russel M. Barefield, "The Effect of Aggregation on Decision Making Success: A Laboratory Study," Journal of Accounting Research (Autumn 1972), pp. 229-242.

eight masters students in Purdue's Master of Science in Industrial Administration program were randomly assigned to two experimental groups. One group received eighty reports containing one cue, an "adjusted labor efficiency" variance (LEV_{adj}), and the other, eighty reports containing two variances, a labor efficiency variance (LEV) and a material quantity variance (MQV). The labor efficiency variance (LEV) and the material quantity variance (MQV) could be mathematically combined to generate the adjusted labor efficiency variance (LEV_{adj}). The two sets of eighty reports were generated using bivariate normal distributions so that the environmental predictablity for each set of reports was the same.

The participants' performance was measured as the difference between the optimal number of correct responses, based on the selections of the environmental models, and the actual number of correct decisions make by the participants. The differences in prediction accuracy were 2.58 for the aggregated data and 1.18 for the disaggregated data. The prediction accuracy of both groups was high, with the participants receiving disaggregated data performing slightly better. The results, however, were not statistically significant.

The participants' ability to select the proper decision criterion (an optimal model) was determined by measuring the differences between the number of optimal correct decisions, the correct decisions generated by the environmental model, less the number of correct decisions generated by the participants' models. The participants receiving aggregated data did slightly better at selecting a model (the average subject difference was .33) than the participants receiving disaggregated data (the average difference was 1.37). The differences were not statistically significant.

The participants' ability to use their own model in a consistent fashion was determined by measuring the differences between the number of correct decisions generated by the participants' models less the number of correct decisions generated by the participants. The participants receiving disaggregated data applied their own models more consistently (the average subject difference was -.19) than the participants receiving aggregated data (the average difference was 2.25). Again, the differences were not statistically significant.

To summarize, the participants receiving disaggregated data appeared to generate a higher prediction accuracy than those using aggregated data. The participants receiving aggregated data appeared to use their models in a more constent fashion and to select less optimal models. All of the differences reported were small and none of the differences were statistically significant.

Abdel-Khalik explored the effects of different levels of aggregation of financial statement information on the accuracy of commercial lending decisions.[106] This study differs in a key respect from the previous study in that no attempt is made to hold the environmental predictability constant for all levels of aggregation.

Two pairs of matched firms were selected; one firm in each pair had defaulted on a loan contract and one firm had not defaulted. For each firm three levels of aggregation financial statement information were generated. The lowest level of aggregation consisted of the financial data that Beaver considered strong predictors of financial failure.[107] The most detailed presentation of the financial information contained, in addition to the ratio data, information from the firm's annual reports, 10-k reports, and Moody's. The intermediate level of aggregation contained information from the above sources that was considered material. Materiality was defined as any item that affected total assets, total liabilities, or sales by ten percent or more.

Questionnaires were mailed to six hundred bank loan officers, two hundred and seven of whom responded. Each questionnaire contained one of the three levels of aggregated data for one of the matched pairs of firms. The parti-

---

[106] A. Rashad Abdel-Khalik, "The Effect of Aggregating Accounting Reports," pp. 104-138.

[107] Willian H. Beaver, "Financial Ratios and Predictors of Failure," Supplement to <u>Journal of Accounting Research</u> (1966), pp. 71-111.

cipants were requested to perform two types of tasks: to allocate scarce loanable funds between the two firms and to make a subjective estimate of the probability of default for each firm. Specifically, the task variables were:

1. A recommendation of a loan for six months.
2. An estimate of the probability of default on that loan.
3. An estimate of the probability that the firm will be a good credit risk during the following three years.
4. A loan recommendation for a three-year term.
5. An estimate of the probability of default on that loan.[108]

Three criteria were used to measure the quality of the participants' decisions: (1) prediction accuracy, (2) consistency, and (3) consensus. Prediction accuracy was tested by determining if the participants' mean estimates of the probability of default, as measured by the five task variables, were greater for the defaulted firm than for the nondefaulted firm. A one-analysis of variance was used to test the five task variables under the three levels of aggregation. There was no statistically significant difference among the treatments for any of the five task variables.

The firms were classified into defaulted and nondefaulted firms and an analysis of variance performed. The F-ratios of all five variables were greater for the defaulted firms. Statistically significant differences in the

---

[108]Abdel-Khalik, "The Effect of Aggregating Accounting Reports," pp. 110.

means of the decisions were found between aggregation levels one and three and two and three for only the defaulted firms. No statistical differences were found for the nondefaulted firms. "This finding, . . . lends support to the theory that more detail is needed whenever the borrowing firm is a marginal, or bad, risk."[109] Lending decisions are affected by the level of aggregation in the case of high risk firms.

Consistency of the loan officers was measured in three ways:

> the consistency between the loan officers' decisions and the estimates of the probability of default, the consistency between decisions while holding loan officers' risk attitude and experience constant, and the consistency between the probability of default and the probability of being a good credit risk.[110]

It was found that the participants using aggregated data were more consistent in their judgments.

The consensus of the participants' decisions was measured by the dispersion of their judgments. The level of aggregation that had the least variance was considered to have the highest degree of consensus. There was no statistical difference between the variances for the different levels of aggregation. No difference was found between the amount of consensus for the different levels

---

[109] Ibid., p. 119.

[110] Ibid., p. 111.

of aggregation.

Three quality criteria, predictive ability, consistency and consensus were examined to determine if they were affected by the level of aggregation. It was concluded that for the nondefault firms, predictive ability was not affected by the level of aggregation; however, when the firms were marginal, the disaggregated data generated a higher prediction accuracy. Participants using aggregated data were found to be more consistent in their decisions. Finally, no differnce was found between the amount of consensus and the levels of aggregation.

Casey, in a 1980 study, had one hundred and twenty-two bank loan officers predict which of ten actual firms had declaired bankruptcy within a three year period.[111] Five of the firms had become bankrupt. The participants were randomly assigned to three treatment grougps; Group I received a three-year set of six financial ratios for each of the firms, Group II, in addition to the ratios, received Income Statements and Balance Sheets with footnotes, and Group III received, in addition to the information received by Group II, the footnotes to the financial statements. The information was presented to the participants in the form of a questionnaire. The paricipants were requested

[111] Cornelius J. Casey, Jr., "Variation in Accounting Information Load: The Effect on Loan Officers' Predictions of Bankruptcy," The Accounting Review, (January 1980), pp. 36-49.

to record both their predictions and the time required to complete the study. The high prior probability of bankruptcy was not revealed to the participants. They were told, instead, to assume the failure rate of their institutions.

The average prediction accuracy based on ten firms was 4.13 correct predictions for Group I, 5.55 correct predictions for Group II, and 5.27 correct predictions for Group II. The average amount of hours the participants spent on the task was 2.49 for Group I, 3.63 for Group II, and 5.87 for Group III. The author defined information overload as "a decline in user performance due to the assimilation of additional information."[112] Analysis of variance was performed, and both the levels of prediction accuracy and the amount of time spent on the task were statistically significant. Pairwise comparisons were performed on both prediction accuracy and the amount of time required to perform the task. The prediction accuracy of Group II was statistically better than Group I; however, the decrease in prediction accuracy between Group II and Group III was not statistically significant. Group II did not spend significantly more time on the task than did Group I. Group III did spend significantly more time on the task than did Group II.

---

[112]Ibid., p. 37.

The author concludes that, "adding financial statement notes to a smaller information load did not improve the prediction accuracy of corporate bankruptcy and did consume a significantly larger amount of loan officer time to process the data."[113] While there was a decrease in prediction accuracy between Group II and Group III, the contention that information overload was present was not supported statistically. There was, however, a significant cost in time for processing the additional footnote information provided to the Group III participants.

2.5 SUMMARY

In this chapter Brunswick's lens model was described and the components of prediction accuracy presented. A selected survey of significant lens model literature in both psychology and accounting was reviewed. A review of literature pertaining to information overload was also undertaken.

The selected lens model studies were reviewed with the following topics being emphasized: (1) the relative prediction accuracy of human beings and mathematical models, (2) the factors that affect prediction accuracy, and (3) the level of confidence decision makers have in their judgments. The accounting studies reviewed came from the re-

---

[113]Ibid., p. 46.

search areas of bankruptcy prediction and the forecasting of stock price changes. In the bankruptcy prediction studies, prediction accuracy was measured by the number of correct responses, and discriminant models were used to capture the prediction process. In stock price forecasting, prediction accuracy was measured using correlation, and the models were captured using multiple linear regression.

2.5.1 Lens Model Studies

The prediction accuracy of both mathematical models and human decision makers was quite low. The highest prediction accuracy was achieved by the environmental models, followed by the composite judge and the judge's own model. There was no statistical difference between the prediction accuracy of the composte judge and that of the participants' own models. The human decision maker was the least accurate decision maker. An exception to these results was the 1976 Libby study in which he found that thiry-three participants performed as well as or better than their models, and in only ten instances did the model outperform the participant. Wright attributed the results to a well defined distal variable, the participants' expertise at the task, and the nature of the data.

The low levels of prediction accuracy achieved by the participants were a result of low environmental predictability and a high degree of random and systematic error in the participants' processing of the cues. The research findings also indicate that there is no significant relationship between prediction accuracy and confidence and that linear models were used by the participants.

An extremely important study was that by Abdel-Khalik and El-Sheshai in which the participants selected their own cues. There was a significant increase in prediction accuracy as a result of using a mathematical model to select the cues.

2.5.2 Information Overload Studies

The review by Miller provides insight into man's ability to analyze data. Miller concluded that the human decision maker can differentiate approximately seven categories within a cue and that approximately seven to ten cues can be integrated into the human decision process. Other researchers determined that prediction accuracy was not only determined by the number of cues but by the relative information content of the cues, the order in which the cues were presented to the decision maker, and the degree of correlation among the cues.

The results of accounting research in the area of information overload were mixed. Einhorn found that participants using disaggregated data generated higher prediction accuracy than those using aggregated data. Abdel-Khalik, in his bankruptcy study, found that when the firms were of high risk the disaggregated data generated a higher prediction accuracy. Casey, on the other hand, found evidence of information overload.

Chapter 3

RESEARCH QUESTIONS AND HYPOTHESES

The purposes of the research, the research questions, and the hypotheses used to answer the research questions are presented in this chapter. The research is an empirical study based on Eyon Brunswick's "lens" model and is designed to determine the impact of both the number of information cues and the method of selection of the cues on prediction accuracy. Prediction accuracy is defined, within the context of the lens model, as the correlation between the prediction of the human decision maker or a mathematical model, as the case may be, and the event being predicted. The event being predicted is known to the researcher, and the selection of the cues can be made by either the human decision maker or by a mathematical model, which is, for the purposes of this study, a multiple linear regression equation.

There are two main sections in this chapter. In the first, the purposes of the research are described. In the second, the research questions are detailed and the hypotheses used to answer the research questions delineated. The research methodology employed to answer the research questions and the statistical tests used to test the hypotheses are advanced in Chapter Four.

3.1 PURPOSES OF THE RESEARCH

There are three purposes of this research. The first is to replicate traditional lens model studies. The other two, more significant purposes are to gain some insight into the amount of increase in prediction accuracy when a model replaces the human decision maker for the selection of cues, and to examine the information overload phenomenon to determine its effects on prediction accuracy.

3.1.1 Cue Usage

The review of human information processing literature in Chapter Two shows that there has been a great deal of research using Brunswick's lens model to capture the decision maker's model and to measure prediction accuracy.[114] In traditional lens model research when prediction accuracy is examined, the decision maker is provided with a predetermined set of cues and is requested to predict an event that is known to the researcher.[115] The decision maker's model is then captured using multiple linear regression, and his prediction accuracy is measured using correlation statistics. The prediction accuracy achieved by the decision maker is, then, a measure of his ability to use

---

[114] Eyon Brunswick, The Conceptual Framework of Psychology.

[115] Traditional lens model research as it relates to prediction accuracy is discussed by Robert Ashton. Robert H. Ashton, Human Information Processing.

the cues provided. The results of traditional lens model research have been with few exceptions consistent: the prediction accuracy of the regression model is greater than that of the human decision maker. This traditional lens model research will be replicated in the current study.

3.1.2 Cue Selection

There is, however, a second major component of prediction accuracy, information choice. Actual prediction situations require not only the use of cues, but also the selection of cues. The cues can be selected either by the human decision maker or by a mathematical process such as regression. According to lens model theory, high prediction accuracy will be achieved when the cues used to make the prediction are highly correlated with the event being predicted.[116] It would therefore be expected that a multiple regression equation would achieve a more optimal cue selection, and a higher prediction accuracy, than the human decision maker. The first major purpose of this research is to gain some insight into the amount of increased prediction accuracy that can be expected to be achieved when a model replaces the human decision maker for the selection of the cues.

3.1.3 Information Overload

Information overload is defined as the point in the

---

[116]Ibid., p. 30.

prediction process at which the addition of new information decreases prediction accuracy. The review of the literature concerning information overload indicates that human decision makers are limited in the number of information cues that they can use effectively. A ceiling on prediction accuracy achieved by the human decision maker is sometimes reached after the examination of a small number of information cues. The addition of new cues often causes prediction accuracy to remain constant or decrease. Slovic and Lichtenstein state that "there is a small amount of evidence that increasing the amount of information available increases his (the decision maker's) confidence without increasing the quality of his decisions."[117] A second major purpose of this research is to examine the information overload phenomenon to provide some research evidence concerning the point at which additional information may contaminate prediction accuracy.

3.2 RESEARCH QUESTIONS

The hypotheses were developed to answer the following research questions:

1. Do the results of traditional lens model research hold for the current study? The results of tradi-

---

[117] Paul Slovic and Sarah Lichtenstein, "Comparison of Bayesian and Regression Approaches to the Study of Information Processing Judgment," Organizational Behavior and Human Performance, (1971), p. 687.

tional lens model research indicate that human decision makers will be out performed by both the environmental model and their own models and that the environmental model will outperform both the human decision maker and the decision maker's model.[118]

2. Will forecasts based on the mathematical selection of the cues be more accurate than forecasts based on the human selection of cues?
3. Does information overload take place as the human decision maker is provided with an increasing number of cues?

Table 3, p. 88, defines the correlation statistics to be used in the testing of the hypotheses generated to answer the research questions one and two.

3.3 HYPOTHESES

Several hypotheses were generated to answer the research questions. Two of the hypotheses relate to the first research question concerning the effect of cue usage on prediction accuracy. Three additional hypotheses relate to the second research question concerning the effect of cue selection on prediction accuracy. The last two hypotheses address the information overload phenomenon.

3.3.1 Cue Use

---

[118]The results of traditional lens model research are discussed in detail in Chapter Two.

Table 3

CORRELATION STATISTICS

FOR THE

MEASUREMENT OF PREDICTION ACCURACY

Decision Maker	Method of Cue Selection: Human	Method of Cue Selection: Mathematical
Human	r_{ah}	r_{am}
Decision maker's model	r_{mh}	r_{mm}
Environmental model	R_{eh}	R_{em}

Definitions:

r_{ah} - the correlation between the distal variable and the decision maker's response using the decision maker's selection of cues.

r_{am} - the correlation between the distal variable and the decision maker's response using mathematically selected cues.

r_{mh} - the correlation between the distal variable and the decision maker's mathematical model using the decision maker's selection of cues.

r_{mm} - the correlation between the distal variable and the decision maker's mathematical model using mathematically selected cues.

R_{eh} - the multiple correlation coefficient between the cues selected by the decision maker and the distal variable.

R_{em} - the multiple correlation coefficient between the mathematically selected cues and the distal variable.

First to be examined will be the change in prediction accuracy that is expected to take place when the decision maker is removed from the prediction process and replaced by a mathematical model. The general hypothesis is that prediction accuracy will increase when the participant is replaced by the mathematical model because of the reduction in random and systematic error. Stated in the alternate form the hypotheses are:

Ha_1: $R_{eh} > r_{mh} > r_{ah}$

Ha_2: $R_{em} > r_{mm} > r_{ah}$

3.3.2 Cue Selection

The effects of cue selection on prediction accuracy will be measured by comparing the prediction accuracy achieved when the decision makers selected their own cues with the prediction accuracy achieved when the cues were provided. Three hypotheses are tested. Stated in the alternate form the hypotheses are:

Ha_3: The prediction accuracy of the decision makers' environmental models with cues provided, R_{em}, is greater than the prediction accuracy achieved when the decision makers select their own cues, R_{eh}. $R_{em}-R_{eh} > 0$.

Ha_4: The prediction accuracy of the decision makers' models with the cues provided, r_{mm}, is greater than the prediction accuracy achieved when the decision makers selected their own cues, r_{mh}.

$r_{mm} - r_{mh} > 0$.

Ha_5: The prediction accuracy of the decision makers with the cues provided, r_{am}, is greater than the prediction accuracy achieved by the decision makers when they selected their own cues, r_{ah}.

$r_{am} - r_{ah} > 0$.

3.3.3 Information Overload

Statistically, the addition of valid information cues increases prediction accuracy; however, researchers have found that the human decision maker may reach a ceiling of prediction accuracy after examining very few cues.[119] After reaching the ceiling, prediction accuracy remains constant or decreases with the addition of new, valid cues.[120] This phenomenon is called information overload.

As decision makers review the data in an attempt to draw conclusions, their level of confidence often increase as the amount of available information increases. Because of information overload, the increase in confidence is often not justified.[121]

---

[119] Linda W. Dudycha and James C. Naylor, "Characteristics of the Human Inference Process in Complex Choice Behavior Situations," Organizational Behavior and Human Performance, (September 1966), pp. 110-29.

[120] Paul Slovic and Sarah Litchtenstein, (1971), p. 687.

[121] Stewart Oscamp, "Overconfidence in Case-study Judgments," Journal of Consulting Psychology, (June 1965), p. 264.

It is hypothesized that prediction accuracy will be maximized after the subject has examined a small number of cues, and a decline in prediction accuracy will then take place as new cues continue to be examined. It is also hypothesized that the decision maker's confidence level will increase as long as new cues are received. Confidence level would therefore not be expected to be related to prediction accuracy. It is therefore hypothesized that:

Ha_6: Prediction accuracy (r_{am}) will be maximized after the decision maker's examination of a small number of cues and then decline as additional cues are examined.

Ha_7: The decision maker's confidence in his prediction accuracy will increase with the use of additional cues.

3.4 SUMMARY

There are two major purposes of this empirical study. The first, is to examine the impact on prediction accuracy of the participants' selecting their own environmental cues. Theoretically, prediction accuracy should decrease because the cues selected by the participants will not be as highly correlated to the environmental event as will the cues selected using a multiple linear regression equation. The second major purpose of the research is to examine the information overload phenomenon to provide research evidence as to the point at which additional information may contaminate prediction accuracy.

In accordance with the stated purposes of this study, research questions dealing with the effects on prediction accuracy of the selection, use and number of cues, have been framed; and hypotheses have been formulated to answer these questions. The study is an empirical study based on Brunswick lens model methodology. The lens model paradigm uses multiple linear regression equations to model the relationship that exists between the human decision maker and the environmental event being predicted. The lens model provides the methodology for measuring prediction accuracy and the theoretical basis for the examination of the factors that affect prediction accuracy.

Chapter 4

METHODOLOGY

This chapter contains the methodology for an experimental study designed to provide answers to the research questions introduced in Chapter Three. The experiments will be briefly described, and the development of the cases employed in the experiment, the verification procedures used to confirm the accuracy of the case data, and the criteria developed to select the participants are discussed. Finally, the statistical procedures utilized to develop the environmental regression models and to analyze the results of the experiments are reviewed.

4.1 DESCRIPTION OF THE EXPERIMENTS

Two experiments were performed. The first examined the impact of cue choice and cue use on prediction accuracy, and the second examined the impact of the number of cues on prediction accuracy. The experiments required the participants to complete one or two sets of cases that were generated from a data base and displayed at a computer terminal.

At the conclusion of the experiments, each participant completed a short questionnaire which elicited basic demographic information, education level achieved, and the level of involvement and interest in the experiments. For the information overload portion of the study, the questionnaire also attempted to determine if the participants developed a prediction strategy, and if they had an intuitive understanding of the number of cues required to maximize prediction accuracy.

4.1.1 Cue Choice and Cue Use

In the first experiment the participant decision makers were provided with two groups of thirty cases each, each case consisting of a set of cues in the form of financial ratios for an actual manufacturing firm. To complete the first set of cases, the Group I case set, the participants were requested to select, from a "shopping list" of thirty ratios, six that they believed to be the best predictors of a manufacturing firm's subsequent year's ROA. The ratios selected were then automatically generated by the computer for each of the thirty firms and displayed in case format at the participants' terminals. An example of a case is provided in Table 4, p. 95.

For each individual case the participant decision makers recorded their predictions of the firm's subsequent year's ROA and their levels of confidence in their predictions. After the completion of each case, the next was automatically displayed at the participants' terminals until all thirty cases had been analyzed.

After the participants finished the Group I case set,

Table 4

EXAMPLE OF A CASE

Case 10

No.	Cue	%
4	Operating income margin for 19x4.	3.4
9	Sales/operating assets for 19x4.	110.3
14	ROA for 19x4.	12.7
20	Sales growth: 19x2 through 19x4.	37.3
26	Industry sales growth for 19x4.	15.9
29	Industry ROA for 19x4.	5.8

Prediction:

Please estimate the expected ROA at 12/31/x5. ____

Confidence level:

Indicate the level of confidence in your judgment. ____

they were requested to complete the second set of thirty cases, the Group II cases. The procedure for the Group II cases was the same as for the Group I cases except that a set of six mathematically selected cues were provided for the participants.

4.1.2 Information Overload

All participants in the second experiment were provided with a "shopping list" of eight financial ratios and requested to select the two ratios they believed to be the best predictors of a firm's subsequent year's ROA. The two ratios were then automatically generated for each of the thirty firms and displayed at the participants' computer terminals in case format. For each case the participants recorded their predictions of the subsequent year's ROA and their levels of confidence in the predictions. Upon completion of the thirty cases, the participants were requested to select two additional cues and repeat the prediction process. For each of the thirty cases, the second set of cues was displayed in addition to the first set, providing further information upon which the decision makers could base their predictions and confidence levels. The process was repeated until all eight cues were used and the decision makers had revised their predictions and confidence levels three times.

4.2 DEVELOPMENT OF THE CASES

The ninety cases required for the experiments were generated using actual financial data from manufacturing firms. Each case consisted of a set of financial ratios that were used by the decision makers as cues to predict the subsequent year's ROA for the firm. The ninety cases were randomly divided into three groups of thirty cases each. Group I cases and Group II cases were used to examine the impact of information choice and information use on prediction accuracy, and Group III cases were used to examine the information overload phenomenon.

4.2.1 Selection of the Firms

The criteria for the selection of the firms included in the cases were as follows:

1. The firm must be a manufacturing company.
2. The firm must be in the COMPUSTAT files, registered on the New York Stock Exchange, and listed in the Primary Industrial File or the Supplementary Industrial File of the Standard and Poor's Industrial Index.
3. Seventy-five percent of the firm's products must be directed towards one specific market.
4. Seventy-five percent of the firm's production must take place within the United States.
5. Seventy-five percent of the firm's sales must be within the United States.
6. No more than twenty-five percent of the firm's sales can be to government agencies.

The criteria were selected to assure that the data required for the creation of the financial ratios would be available and that the ratios given to the decision makers for use as cues would provide insight into the factors that affect the firm's subsequent year's ROA.

The selection of the ninety firms to be adopted as cases for the experiments was a four step process. Initially, a list of firms was compiled by selecting from the COMPUSTAT tapes those firms listed on the New York Stock Exchange and contained in either the Primary Industrial File or the Supplementary Industrial File of the Standard and Poor's Industrial Index. Then, the published financial statements for each of the firms on the initial list were examined; the firms that did not meet all six criteria were eliminated. Next, if the data required for the generation of a firm's financial ratios was either not available or did not appear capable of being verified, those firms were eliminated. The above process generated a list of ninety-five firms. Five firms which appeared to least fit the six criteria were eliminated, and the resulting ninety firms were used to create the case sets.

4.2.2 Selection of the Cues

Research has found that macroeconomic, industry, and company factors all interact to affect a firm's ROA in both the short and long term and that many of these factors are effectively captured in the form of financial ratios.[122] The cues used to generate the cases are, therefore, financial ratios created using both individual firm and industry data. The individual firm data was generated from the COMPUSTAT tapes and the industry data from Value Line.

A list of potential cues was generated from two sources. Accounting and financial research literature was examined to find financial ratios that research indicates are related to or are predictors of a firm's ROA. Financial services for investors were reviewed to determine the financial ratios that were being published for use by actual decision makers. There were two criteria for the use of a ratio as a cue; the ratio or the data needed to compute the ratio had to be available in either the COMPUSTAT tapes or in published financial services, and the ratios or the data comprising the ratios had to be capable of verification.

4.2.3 Selection of the Participants

The participants for the study were selected from students enrolled in the Masters of Accountancy and the Masters in Business Administration programs at Virginia Polytechnic Institute and State University. Students selected to participate in the study were expected to have

---

[122] Joel G. Siegel, "How Management Can Make Significant Use of Financial Analysis to Evaluate the Health of A Firm," The National Public Accountant, (February 1982), pp. 25-9.

an educational background that would assure that they understood the prediction process and were familiar with the factors affecting ROA. The minimum requirement for the Masters of Accountancy student was the completion of the required undergraduate courses in intermediate and cost accounting for entry into the masters program. The minimum requirement for selection of MBA students was the completion of all of the prerequisite courses in accounting and finance and the completion of either the required introductory graduate finance course or the required introductory graduate level accounting course.

4.2.4 Development of the Questionnaire

At the conclusion of the experiment, each participant completed a short questionnaire. The purpose of the questionnaire was to elicit basic demographic information, the educational level achieved by the participant, information about the participant's career interests, and the level of involvement and interest in the cases. For the information overload portion of the study, the questionnaire also attempted to ascertain if the subjects developed prediction strategies and if they had an intuitive understanding of the number of cues required to maximize prediction accuracy.

4.2.4 Pilot Study

Before the actual study was attempted, each member of the author's dissertation committee performed the experiments; and a pilot study was performed using as participants ten Ph. D. students in accounting and finance. Several factors were revealed as a result of this preliminary process. The logon procedure was found to be time consuming because the passwords were not clear, and the subjects could not determine the difference between a zero (0) and the letter O in the logon password. There was also a consensus that the experiment was boring and that the information overload portion was too long. As a result of these findings, changes were made in the experimental procedure, the experiment itself, and the questionnaire. The questionnaire was expanded to determine if the participant found the experiment boring or fatiquing, and if so, at what point in the decision making process the boredom or fatigue began. The number of cues required to complete the information overload portion of the study was reduced from ten to eight in an attempt to reduce the amount of time needed to complete the experiment and to lessen the fatigue and boredom factors. Finally, the participant was not required to logon to the system in order to begin the experiment. The experimenter completed the logon process for all the participants before the beginning of the experiment.

4.3 PROCEDURES FOR VERIFYING THE DATA

A data base consisting of thirty-one financial ratios for each of the ninety firms selected to be used as cases was created by drawing from two sources. The COMPUSTAT tapes provided financial statement information pertaining to the individual firms, and *Value Line Investment Survey: Ratings and Reports* provided industry data. Thirty of the ratios were independent variables used to predict the dependent variable, subsequent year's ROA.

Information from four consecutive years was required for the creation of the ratios. The data was selected from the years 1979 through 1982. The 1979, 1980, and 1981 data was utilized in the generation of the independent variables. The 1982 data was employed in the creation of the dependent variable, and the firm's ROA for that year.

4.3.1 Verification of Individual Firm Ratios

Because *Moody's Industrial Manuals* contained the requisite information, it was used as the primary source for the verification of the firm data extracted from the COMPUSTAT tapes. To confirm the accuracy of *Moody's*, fifteen published financial statements, for the years 1979 through 1982, were randomly selected from the ninety firms used as cases, and traced to the financial statements reported in the *Moody's Industrial Manuals*. No differences were found.

After the accuracy of *Moody's* was established, the data used to create the ratios was verified by randomly selecting thirty firms from the ninety used as cases, and tracing the data extracted from the COMPUSTAT tapes to the firm's financial statements reported in Moody's. Both the year and the amount of the data item were verified. With the exception of research and development, and advertising expenses, which could not be verified from the published data, no material errors were found. A material error was defined as a ten percent difference between the amount of the data item reported in Moody's and the amount of the data item drawn from the COMPUSTAT tapes.

When the regression was performed to select the environmental models for the Group II and Group III case sets, it was determined that the ratios requiring the research and development and advertising expense information were not predictors of future ROA. Consequently, those ratios were deleted from the cases and no further attempt was made to verify this data.

After the data was extracted from the COMPUSTAT tapes, a program was written to convert the data into ratios. Then, twenty firms were randomly selected and all the ratios for each of the twenty firms were manually recomputed using the raw data generated from the COMPUSTAT tapes. No discrepancies were found.

4.3.2 Verification of Industry Ratios

The raw data for the industry ratios contained in the cases was extracted from Value Line Investment Services, Inc. All of the data was retraced and recomputed to confirm the accuracy of the transcription of the data and the calculations performed.

4.3.3 Verification of the Transcription of the Ratios into Data Bases

After completing the verification procedures described above, the data base consisting of the thirty-one financial ratios for each of the ninety firms was randomly divided into three data bases of thirty firms each from which the Group I, Group II, and Group III cases were generated. The Group I case set required all thirty-one of the cues generated for the thirty cases, the Group II cases required seven cues, and the Group III cases required nine cues. The cues for the Group II and the Group III cases were selected using multiple linear regression. The correctness of the transcription of the data into the three data bases was confirmed by retracing the data from the original data base to the three individual data bases. No errors were found.

4.3.4 Verification of the Programs Creating the Cases

The programs created for generating the cases from the individual data bases were verified by tracing the data displayed on the computer terminal to the original data base. No errors were found.

4.4 CREATION OF THE REGRESSION MODELS

Two types of multiple linear regression equations were created for the purposes of this study. The environmental models, representing the "optimal" selection and weighting of the cues, and the decision makers' models, containing the cues selected by the participant decision makers and their weighting of those cues, were generated.

4.4.1 The Decision Makers' Models

Theoretically, the decision makers' models capture the participants' decision processes, and the coefficients of multiple correlation of the models serve as a measures of the participants' prediction accuracy. The decision makers' models were created by regressing the participants' decisions against the cues upon which the decisions were based.

4.4.2 The Environmental Models

The environmental regression models served two purposes in this study. First, the creation of the models generated the cues that were provided for the participant decision makers in the Group II and Group III case sets. Second, the multiple correlation coefficients of the environmental models indicated the upper level of prediction accuracy that could be expected to be achieved by the decision makers.

The coefficient of multiple determination, R^2, indicated the degree of the linear association between the dependent variable, subsequent year's ROA, and the independent variables, the financial ratios. The square root of the coefficient of multiple determination is the multiple correlation coefficient, R, the measure of prediction accuracy for the purposes of this study and the benchmark against which the prediction accuracy of both the human decision makers and their models are compared. The primary goal in the creation of the environmental models, therefore, was to select from the "shopping list" of thirty financial ratios, the independent variables that appeared to maximize the coefficient of multiple determination.

A secondary goal, not to be achieved at the expense of a material reduction in the coefficient of multiple determination, was to select models in which the independent variables were not highly correlated. The presence of multicollinearity has two potentially serious effects on the individual regression coefficients. First, the variances of the coefficients will be large, indicating that the coefficients may be poorly estimated and that the estimates are unstable.[123] Second, multicollinearity tends to inflate the regression coefficients. For the environmental equations to be used as models against which the participants' models could be compared to gain insight to the effects of cue selection and cue weighting on prediction accuracy, the individual regression coefficients have

[123]Montgomery and Peck, pg 391-2.

to be accurately determined. Multicollinearity had therefore to be controlled.

The number of cues contained in the environmental models was restricted because regression models should include only the independent variables that make a significant contribution to the prediction accuracy of the model. Montgomery and Peck explain. "Building a regression model that includes only a subset of the available regressors involves two conflicting objectives. (1) We would like the model to include as many regressors as possible so that the 'information content' in these factors can influence the predicted value of y. (2) We want to include as few regressors as possible because the variance of the prediction y increases as the number of regressors increases."[124] The increased variance of y, the fitted regression line, causes the environmental models to be less reliable for the purpose of prediction. The number of cues used by the decision makers to analyze the Group I and Group II cases sets was limited to six, and the number of cues used by the decision makers to analyze the Group III case sets was limited to eight.

4.4.2.1 Creation of the Environmental Models

The creation of the environmental regression models was a two step process. First, a number of models were

---

[124] Ibid., p. 244.

generated using the t-statistics to select the cues that appeared to contribute significantly to the coefficient of multiple determination. Second, the variance inflation factors and condition indexes for each of the models generated were examined, and the models for which collinearity between the independent variables was controlled and for which the cues appeared to be logical predictors of the dependent variable were selected.

The t-statistics were used to select the independent variables for the environmental models because of the large number of independent variables from which to choose. Montgomery and Peck explain that the use of the t-statistic "is often a very effective variable selection strategy when the number of candidate regressors is relatively large, for example K > 20 or 30."[125]

4.4.2.2 Detection of Multicollinearity

Variance inflation factors and condition indexes were used to detect the presence of multicollinearity and to determine the amount of collinearity between cues that would be detrimental to the results of this study. Variance inflation factors (VIFs) are created using the inverse of the correlation matrix of the explanatory variables. The amount of VIF that indicates that multicollinearity may cause the individual regressors to be unreliable has

---

[125] Ibid., p. 266.

been defined by Montgomery and Peck. "Practical experience indicates that if any of the VIFs exceeds 5 or 10, it is an indication that the associated regression coefficients are poorly estimated because of multicollineaity."[126] Models which contained variance inflation factors of greater than 5 were eliminated from consideration.

Condition indexes are generated from the eigenvalues of the correlation matrix of explanatory variables. The advantage of the eigensystem is that it detects collinearity between two or more explanatory variables and that indexes for the determination of multicollinearity and the variables affected have been determined from the "extremely rich literature in numerical analysis."[127]

A small eigenvalue indicates the presence of multicollinearity. Belsley, Kuhn, and Welsh recommend the use of condition indexes to determine the meaning of small and the use of regression-coefficient variance decomposition proportions to determine which coefficients are degraded because of multicollinearity. The condition index is the smallest eigenvalue divided by the largest eigenvalue. Belsley, Kuhn, and Welsch state that "weak dependencies are associated with condition indexes around 5 or 10, whereas moderately strong relations are associated with indexes

---

[126]Ibid., p. 300.

[127]Belsley, Kun, and Welsh, p. 96.

of 30 to 100."[128]

The variance-decomposition proportions are also based on eigenvalues and are the proportions of variance shared between the variables. Variance-decomposition proportions of over 0.5 for two coefficients indicate collinearity between the coefficients. Multicollinearity would be indicated by a condition index of greater than 30 associated with variance-decomposition proportions of greater than 0.5. Models that had a condition index of over 20 and a variance-decomposition proportion of over 0.5 were eliminated from consideration.

The model used in this study was chosen from those remaining by selecting what appeared to be the most logical model. The independent variables appeared to be related to the dependent variable and the signs on the coefficient appeared rational.

To select the final environmental models to be used for the Group II and Group III case sets diagnostic statistics for the detection of multicollinearity were examined for each model. The final models selected for each of the case sets had the following properties:

1. A high R^2. No material difference existed between R^2 of the model chosen and the models not chosen.
2. Multicollinearity was controlled.

---

[128]Ibid., p. 105.

3. The independent variables appeared to be rational; that is, they all appeared to be related to the dependent variable in a logical fashion.

It should be noted, however, that while the presence of strong multicollinearity produces poor estimates of the individual regression coefficients, the multiple correlation coefficient of the environmental models, the measure of prediction accuracy used in this study, is not affected by the amount of multicollinearity between the independent variables.

4.5 STATISTICAL ANALYSES EMPLOYED

The procedures employed to select the statistical tests applied to the hypotheses presented in Chapter Three, and the descriptive statistics used to delineate the relationships contained in the data are discussed in this section. In addition, the statistical procedures employed to summarize the results of the post test questionnaire are reviewed.

4.5.1 Cue Selection

To control for the differences in the predictive abilities of the individual participants, paired observations were used; each participant first analyzed the Group I case set and then the Group II case set. There are three potential statistical tests available for paired data; the sign test, the Wilcoxon signed ranks test, and the paired-t test. The paired-t test is a parametric procedure and the most powerful of the tests; however, three assumptions must be satisfied for its use:

1. The distribution of the differences in prediction accuracy ($R_{eh}-R_{em}$, $r_{mh}-r_{mm}$, and $r_{ah}-r_{am}$) is symmetrical.
2. The distrubituions of the two populations (R_{eh} and R_{em}, r_{mh} and r_{mm}, and r_{ah} and r_{am}) are equally dispersed; they have equal variances.
3. The distribution of the differences in prediction accuracy ($R_{eh}-R_{em}$, $r_{mh}-r_{mm}$, and $r_{ah}-r_{am}$) is normal.

The Wilcoxon signed rank test is a nonparametric analysis that requires the satisfaction of the first two of the above assumptions. The sign test, also a nonparametric test, is the least powerful procedure, but none of the assumptions need to be satisfied for its use. Visual representations of the data and statistical tests of the individual assumptions were used to select the appropriate test.

4.5.2 Cue Usage

To test the hypotheses associated with cue usage, a randomized complete block ANOVA design with multiple comparisons and ordered alternatives procedures was employed. It was assumed that the participants in the study were not homogeneous with regard to their abilities to predict future ROA, and the differences among the participants' predictive abilities were controlled for by using a randomized complete block design. The procedure is a two-way ANOVA that measures the effects of two factors, the treatments (the differences in the prediction accuracies of the environmental models, the participants, and the participants' models) and the blocks (the variation in prediction accuracy among the participants). The blocking, by controlling for differences in prediction accuracy among the participants, makes it more likely that the variation in prediction accuracy caused by the treatments will be detected. The randomized complete block design also provides F-statistics for the blocks, as well as for the treatments, providing a test for the assumption that the participants are not homogeneous in their predictive abilities.

Two critical assumptions required for the utilization of parametric analysis of variance procedures are:

1. The observations are normally distributed for each population.
2. The variance of the observations is the same in each population.

Although the ANOVA is a robust procedure, one in which moderate departures from the above assumptions do not materially affect the results of the analysis, the data from the Group II case set was obviously a significant violation of both assumptions. The prediction accuracy of the environmental models of all thirty-one participants for the Group II case set was $R_{em} = 0.7787$. Nonparametric procedures were therefore appropriate for the analysis of the Group II case set. To be consistent, nonparametric procedures were also used to analyze the Group I case set.

4.5.3 Information Overload

The above discussion of the randomized complete block ANOVA model and the assumptions for parametric procedures are equally applicable for the tests of the hypotheses pertaining to information overload. Again, the participants were blocked because the differences between the prediction accuracy of the individual participants comprise the factor being controlled. The examiantion of the information overload phenomenon, however, required two tests, the first to determine the effect of the number of cues on prediction accuracy, and the second to determine the effect of the number of cues on the participant's confidence level. The treatments for both tests were the number of cues provided to the decision makers.

As stated above, strict compliance to the assumptions of the model is not required. Kleinbaum and Kupper state that while the ANOVA is a robust procedure,

> We must nevertheless be careful to avoid using robustness as an automatic justification for blindly applying the ANOVA method. . . .the normality assumption does not have to be exactly satisfied as long as we are dealing with relatively large samples (e.g., 20 or more observations from each population), although the consequences of large deviations from normality are somewhat more severe for random factors than for fixed factors. The assumption of variance homogeneity can also be violated without serious risk, pro-

> vided that the numbers of observations selected from each population are more or less the same . . .[129]

There are no good tests of the assumptions required for the parametric procedures. The determination of the appropriateness of parametric procedures must be made through visual examination of the data. Inman and Conover explain,

> good tests for normality essentially do not exist for the experimental designs about to be introduced. The experimenter should visually check for obvious signs of nonnormality, such as discrete valued data or the presence of outliers. In the absence of obvious nonnormality, the usual practice is to use the procedures that ... are based on the normality assumption.[130]

To summarize, hypotheses associated with both prediction accuracy and confidence level were tested using a randomized complete block ANOVA design with multiple comparisons and ordered alternatives procedures. Because there are no satisfactory statistical tests for the assumptions of normality and equal variance, the selection of parametric or nonparametric procedures was based on a visual examination of the data.

4.5.4 Analysis of the Post-Test Questionnaire

Several participants in the pilot study reported that

---

[129] David G. Kleinbaum and Lawerence L. Kupper, Applied Regression Analysis and Other Multivariable Methods (North Scituate: Duxbury Press, 1978), p. 248.

[130] Ronald L. Iman and W. J. Conover, Modern Business Statistics (New York: John Wiley & Sons, 1983), p. 555.

the experiment was boring, fatiguing, and unrealistic. These assertions lead to the question, does the participant's perception of the realism of the study and the amount of fatigue experienced by the participant relate to the level of prediction accuracy achieved?

The participant's perceptions of the three factors were captured in the post-test questionnaire by using a seven point Likert scale. To determine if the results of the study were affected by these factors three correlations were examined:

1. The participant's reported level of interest in the study and the prediction accuracy achieved.
2. The level of fatigue reported by the participant and the prediction accuracy achieved.
3. The perceived level of realism of the study reported by the participant and the prediction accuracy achieved.

Because a Likert scale was used to capture the data, the data cannot be assumed to be equal-interval data as required for the use of parametric correlation procedures. As a result, Spearman's coefficient of rank correlation was utilized to analyze the above relationships.

4.6 SUMMARY

In this chapter an experiemental design based on a case study metho dology was reviewed. The experiments, including the procedures for selecting and verifying the data used to generate the cases, were described. The criteria used to develop the environmental regression models and to select the statistical procedures utilized to test the hypotheses presented in Chapter Three were delineated. The results of the tests of hypotheses will be presented in Chapter Five and conclusions suggested by the results will be discussed in Chapter Six.

Chapter 5

RESULTS OF THE STUDY

In this chapter the research methodology presented in Chapter Four is used to test the hypotheses developed to answer the research questions presented in Chapter Three. As in the previous chapters, Chapter Five contains three major sections; cue usage, cue choice, and information overload. Cue choice will be presented first because the description of the variables used to test the cue choice hypotheses are equally apropos to cue choice and information overload. Finally, the results of a post-test questionnaire designed to obtain demographic information about the participants in the study, capture their attitudes towards the study and determine their level of interest in the experiment will be examined.

5.1 INFORMATION CHOICE

As discussed in Chapter Three, a major objective of this study is to determine if forecasts employing environmental cues selected by a mathematical model are more accurate than forecasts based on cues selected by human decision makers. It was expected that the mathematical selection of cues would yield more accurate forecasts because the use of multiple linear regression to select the cues would generate the "optimal" set of cues and weight them so as to maximize the multiple correlation coefficient, the definition of prediction accuracy for the purposes of this study. Three distinct measures of prediction accuracy were used: R_e, the multiple correlation coefficient of the environmental regression models, measures the prediction accuracy of the environmental models; r_a measures the prediction accuracy of the human decision maker; and r_m measures the prediction accuracy of the decision maker's model.

5.1.1 Data

To generate the statistics needed to test the cue choice hypotheses four types of data sets were required for each case set, Case Set I and Case Set II, and the original ninety firms. They are:

1. The environmental event being predicted by the decision maker (Y_e). The environmental event is the return on assets for the year 1982 for each of the cases and was generated from the COMPUSTAT data base.

2. The decision makers' predictions of the 1982 ROA for each of the firms contained in the two case sets (Y_s). The firm's 1982 ROA was presented to the participants as "ROA for year x5" to disguise the time period from which the data used to generate the cases was drawn. The predictions were captured by the computer as the participants completed each case.

3. The optimal prediction of "ROA for year x5" generated

by the environmental regression equations ($\hat{Y}_e$). There are two distinct types of environmental models, the model containing the participants' cues ($\hat{Y}_{eh}$) and the model containing the "optimal" cue set ($\hat{Y}_{em}$). The predictions of the participants' environmental models were obtained using a multiple regression equation containing the participants' cue selections. The predictions of the "optimal" environmental model were obtained using a multiple regression equation containing the "optimal" cue set.

4. The optimal prediction of the participants' responses in prognosticating "ROA for year x5" generated by the participants' models ($\hat{Y}_s$). The participants' models were captured by regressing the predictions of each participant against the cues he selected. This regression model, the decision maker's model, was then used to generate a prediction of the environmental event for each firm contained in each of the case sets.

The notation for the individual data sets is presented in Table 5, p. 121, and the relationships among the data sets are summarized in Figure 1, p. 20.

5.1.2 Statistics

From the above data sets the statistics were generated to test the hypotheses for both the cue choice and cue use portions of the study. Three sets of statistics were generated to measure three distinct aspects of prediction accuracy: R_e, the multiple correlation coefficient of the

Table 5

DATA SETS

Description	Case Sets: I	II	Ninety
Environmental event	Y_{eI}	Y_{eII}	Y_{e90}
Participants' predictions of the environmental event	Y_{sI}	Y_{sII}	
Prediction of the distal variable by the environmental models using:			
The participants' cue sets	$\hat{Y}_{ehI}$		$\hat{Y}_{eh90}$
The "optimal" cue set	$\hat{Y}_{emI}$	$\hat{Y}_{emII}$	$\hat{Y}_{em90}$
Prediction of the participants' responses by the participants' models	$\hat{Y}_{sI}$	$\hat{Y}_{sII}$	

environmental regression models, measures the prediction accuracy of those models; r_a measures the prediction accuracy of the human decision makers; and r_m measures the prediction accuracy of the decision makers' models.

5.1.2.1 Case Set I

In Case Set I the participants selected their own cues. Four statistical measures of prediction accuracy were generated from Case Set I data:

1. R_{ehI} is the measure of the prediction accuracy of the participants' environmental models. R_{ehI} is a multiple correlation coefficient generated by regressing the participants' cue selections against the environmental event. $R_{ehI} = R_{YeI\hat{Y}ehI}$.
2. R_{emI} is the measure of the prediction accuracy of the "optimal" cue set. R_{emI} is the multiple correlation coefficient generated by regressing the cue set of the "optimal" environmental model, as described in Chapter Four, against the environmental event. $R_{emI} = R_{YeI\hat{Y}emI}$.
3. r_{ahI} is the measure of the prediction accuracy of the individual participant. r_{ahI} is the degree of correlation, measured using the Pearson Product Moment Correlation Coefficient, between the environmental event and the participant's prediction of the environmental event. $r_{ahI} = r_{YeIYsI}$.
4. r_{mhI} is the measure of the prediction accuracy of

the decision makers' models. r_{mhI} is the Pearson Product Moment Correlation Coefficient between the participants' models' prediction of the environmental event and the environmental event. $r_{mhI} = r_{YeI\hat{Y}sI}$.

5.1.2.2 Case Set II

In Case Set II the "optimal" set of cues was provided for the participants. Three sets of statistics were generated for Case Set II:

1 .R_{emII} is the measure of the prediction accuracy of the "optimal" cue set. R_{emII} is the multiple correlation coefficient generated by regressing the cue set of the "optimal" environmental model against the environmental event. $R_{emII} = R_{YeII\hat{Y}emII}$.

2. r_{amII} is the measure of the prediction accuracy of the individual participant. r_{amII} is the degree of correlation, measured using the Pearson Product Moment Correlation Coefficient, between the environmental event and the participant's predictions of the environmental event. $r_{amII} = r_{YeIIYsII}$.

3. r_{mmII} is the measure of the prediction accuracy of the decision makers' models. r_{mmII} is the Pearson Product Moment Correlation Coefficient between the participants' models' prediction of the environmental event and the environmental event. $r_{mmII} = r_{YeII\hat{Y}sII}$.

These statistics differ from those in section 5.1.2.1 only in that they are applied to a different set of cases.

5.1.2.3 Case Set of Ninety

The Case Set of Ninety is the original data set for ninety firms before the data set was split into the three case sets. The notation for the Case Set of Ninety will be used in Chapter Six where the adequacy of procedures used to select the "optimal" cue set are discussed. The notation is presented at this point for the sake of clarity.

1. R_{eh90} is the measure of the prediction accuracy of the participants' environmental models. R_{eh90} is a multiple correlation coefficient generated by regressing the participants' cue selections against the environmental event. $R_{eh90} = R_{Ye90\hat{Y}eh90}$.
2. R_{em90} is the measure of the prediction accuracy of the "optimal" cue set. R_{em90} is the multiple correlation coefficient generated by regressing the cue set of the "optimal" environmental model, as described in Chapter Four, against the environmental event. $R_{em90} = R_{Ye90\hat{Y}em90}$.

Table 6, p. 125, summarizes the notation discussed above.

5.1.3 Discussion of the Data

Three hypotheses were tested in conjunction with the cue choice portion of the study:

Ha_3: The prediction accuracy of the decision makers' environmental models with the cues provided, R_{emI}, is greater than the prediction accuracy achieved when the decision makers select their own cues,

Table 6

MEASURES OF PREDICTION ACCURACY

Description	Definition
Case Set I	
Environmental models using:	
The participants' cue sets	$R_{ehI} = R_{YeI\hat{Y}ehI}$
The "optimal" cue set	$R_{emI} = R_{YeI\hat{Y}emI}$
Participant decision makers	$r_{ahI} = r_{YeIYsI}$
Decision makers' models	$r_{mhI} = r_{YeI\hat{Y}sI}$
Case Set II	
Environmental model using the "optimal" cue set	$R_{emII} = R_{YeII\hat{Y}emII}$
Participant decision makers	$r_{amII} = r_{YeIIYsII}$
Decision makers' models	$r_{mmII} = r_{YeII\hat{Y}sII}$
Case Set of Ninety	
Environmental models using:	
The participants' cue sets	$R_{eh90} = R_{Ye90\hat{Y}eh90}$
The "optimal" cue set	$R_{em90} = R_{Ye90\hat{Y}em90}$

R_{ehI}. $R_{emI} - R_{ehI} > 0$.

Ha_4: The prediction accuracy of the decision makers' models with the cues provided, r_{mmII}, is greater than the prediction accuracy achieved when the decision makers selected their own cues, r_{mhI}. $r_{mmII} - r_{mhI} > 0$.

Ha_5: The prediction accuracy of the decision makers with the cues provided, r_{amII}, is greater than the prediction accuracy achieved by the decision makers when they selected their own cues, r_{ahI}. $r_{amII} - r_{ahI} > 0$.

It is expected that the prediction accuracy achieved when the "optimal" set of cues is used will exceed the prediction accuracy achieved when the participants select their own cues. In this section the data for the tests of the hypotheses will be examined, and in the next section the statistical procedures used to test the hypotheses will be discussed.

The first hypothesis uses Case Set I data only. It is expected that the prediction accuracy of the environmental model containing the "optimal" cue set, R_{emI}, will be higher than the prediction accuracy of the environmental models containing the participants' cues, R_{ehI}. Table 7, p. 127, contains a listing of the prediction accuracies achieved by the two environmental models and the differences in the prediction accuracies between the models. The predic-

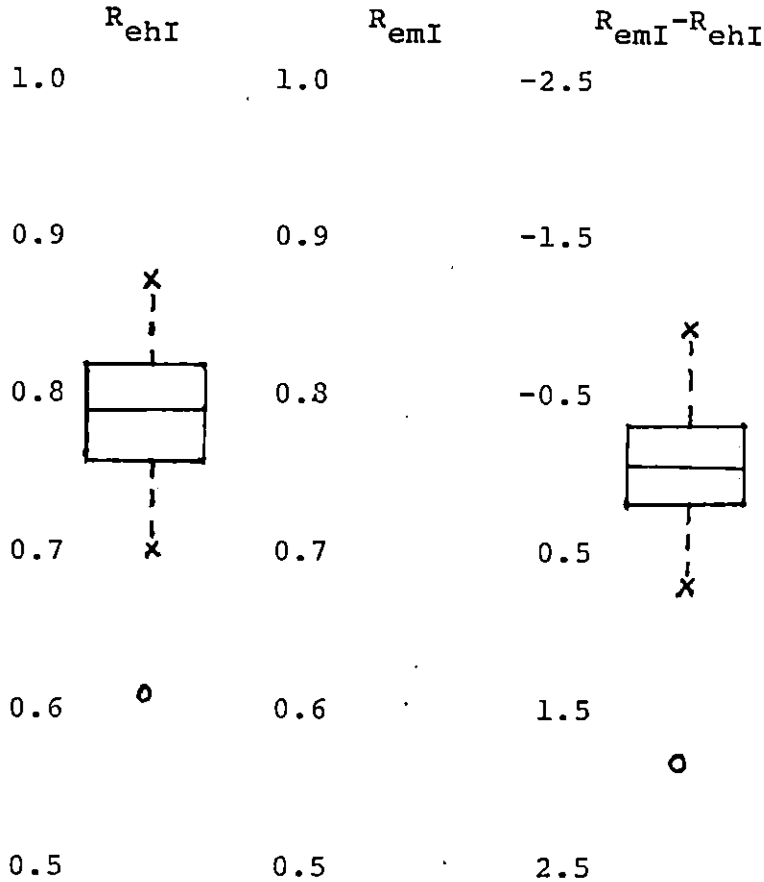
Table 7

PREDICTION ACCURACIES OF THE ENVIRONMENTAL MODELS

Participant	R_{ehI}	R_{emI}	$R_{emI} - R_{ehI}$
14	0.8694	0.7925	-0.0769
30	0.8657	0.7925	-0.0732
18	0.8577	0.7925	-0.0652
24	0.8445	0.7925	-0.0520
2	0.8320	0.7925	-0.0395
4	0.8313	0.7925	-0.0388
29	0.8310	0.7925	-0.0385
11	0.8291	0.7925	-0.0366
15	0.8288	0.7925	-0.0363
25	0.8260	0.7925	-0.0335
1	0.8243	0.7925	-0.0318
21	0.8238	0.7925	-0.0313
17	0.8229	0.7925	-0.0304
10	0.8145	0.7925	-0.0220
13	0.8130	0.7925	-0.0205
19	0.7954	0.7925	-0.0029
12	0.7855	0.7925	0.0070
20	0.7828	0.7925	0.0097
28	0.7781	0.7925	0.0144
7	0.7776	0.7925	0.0149
23	0.7732	0.7925	0.0193
27	0.7730	0.7925	0.0195
8	0.7684	0.7925	0.0241
31	0.7684	0.7925	0.0241
5	0.7597	0.7925	0.0328
22	0.7477	0.7925	0.0448
26	0.7477	0.7925	0.0448
6	0.7469	0.7925	0.0456
16	0.7391	0.7925	0.0534
12	0.7302	0.7925	0.0623
3	0.6117	0.7925	0.1808
Means	0.7934	0.7925	-0.0009
Medians	0.7954	0.7925	-0.0029
Difference in the medians	-0.0029		

tion accuracy achieved by the "optimal" model, R_{emI} = 0.7925, was exceeded by sixteen of the thirty-one participants. The range of the prediction accuracy achieved by the participants was between R_{ehI} = 0.7302 and R_{ehI} = 0.8694, excluding participant number three, who achieved a prediction accuracy of R_{ehI} = 0.6117, which the box plot of the data (Figure 4, p. 129) shows to be a moderate outlier. The mean difference was -0.0009 and the median difference was -0.0029. The box plot also shows the data to be fairly evenly distributed around the median.

Table 8, p. 130, contains the prediction accuracies achieved by the participants when they selected their own cue set, Case Set I, and when the "optimal" cue set was provided to them, Case Set II. The mean prediction accuracy achieved by the participants when they selected their own cues, r_{ahI} = 0.5797, was lower than the prediction accuracy achieved with the cues provided, r_{amII} = 0.6574. The mean difference, 0.0777, was in the hypothesized direction. The difference in the medians, -0.0513, however, was not in the hypothesized direction. Nineteen of the thirty-one participants achieved a higher prediction accuracy when selecting their own cues. The box plot of the data (Figure 5, p. 131) shows that the difference between the mean and median is a result of several extreme outliers, particularly participants number twenty-nine and five. Participant number five, for example, achieved an extremely high predic-



Median	0.7954	0.7925	-0.0029
First Quartile	0.7684		-0.0365
Third Quartile	0.8290		0.0241
Interquartile Rank	0.0606		0.0606
Step	0.0909		0.0909
Inner Fence	0.9199		-0.1274
	0.6765		0.1150
Outer Fence	1.0000		-0.2183
	0.5866		0.2059

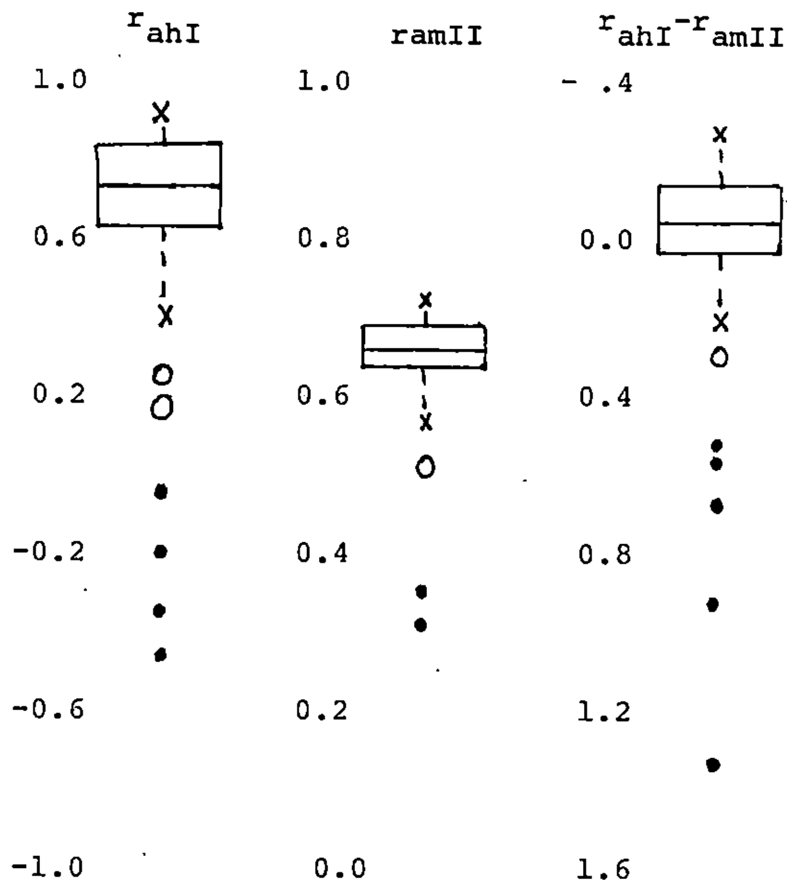
Figure 4

BOX PLOT: PREDICTION ACCURACIES
OF THE ENVIRONMENTAL MODELS

Table 8

PREDICTION ACCURACIES OF THE PARTICIPANTS

Participant	r_{ahI}	r_{amII}	$r_{ahI} - r_{amII}$
17	0.6896	0.2982	-0.3914
24	0.7456	0.5588	-0.1877
8	0.8148	0.6739	-0.1409
1	0.8160	0.6783	-0.1377
6	0.7957	0.6744	-0.1213
27	0.8167	0.6985	-0.1182
16	0.8079	0.6911	-0.1168
22	0.8225	0.7064	-0.1161
15	0.7052	0.5895	-0.1067
10	0.7961	0.6976	-0.0985
13	0.8049	0.7087	-0.0962
18	0.7755	0.6842	-0.0913
7	0.6904	0.6014	-0.0890
31	0.7734	0.6853	-0.0881
12	0.7505	0.6894	-0.0611
26	0.7711	0.7198	-0.0513
23	0.7604	0.7106	-0.0498
30	0.7102	0.6720	-0.0382
11	0.6576	0.6321	-0.0255
19	0.6458	0.6576	0.0118
25	0.6610	0.6761	0.0151
4	0.6998	0.7186	0.0188
20	0.6779	0.7000	0.0221
2	0.6212	0.6650	0.0438
9	0.5022	0.6478	0.1456
3	0.3306	0.6463	0.3157
28	-0.1679	0.3260	0.4939
21	0.1894	0.6903	0.5009
14	0.0772	0.6937	0.6165
29	-0.3386	0.6659	1.0045
5	-0.4329	0.9127	1.3456
Means	0.5797	0.6574	0.0777
Medians	0.7052	0.6783	-0.0513
Difference in the medians	-0.0269		



	r_{ahI}	ramII	$r_{ahI}-r_{amII}$
Median	0.7052	0.6783	-0.0513
First Quartile	0.6335	0.6527	-0.1118
Third Quartile	0.7856	0.6981	0.0330
Interquartile Rank	0.1521	0.0454	0.1448
Step	0.2282	0.0681	0.2172
Inner Fence	1.0000	0.7662	-0.3290
	0.4053	0.5846	0.2502
Outer Fence	1.0000	0.8343	-0.5462
	0.1771	0.5165	0.4674

Figure 5

BOX PLOT: PREDICTION ACCURACIES
OF THE PARTICIPANTS

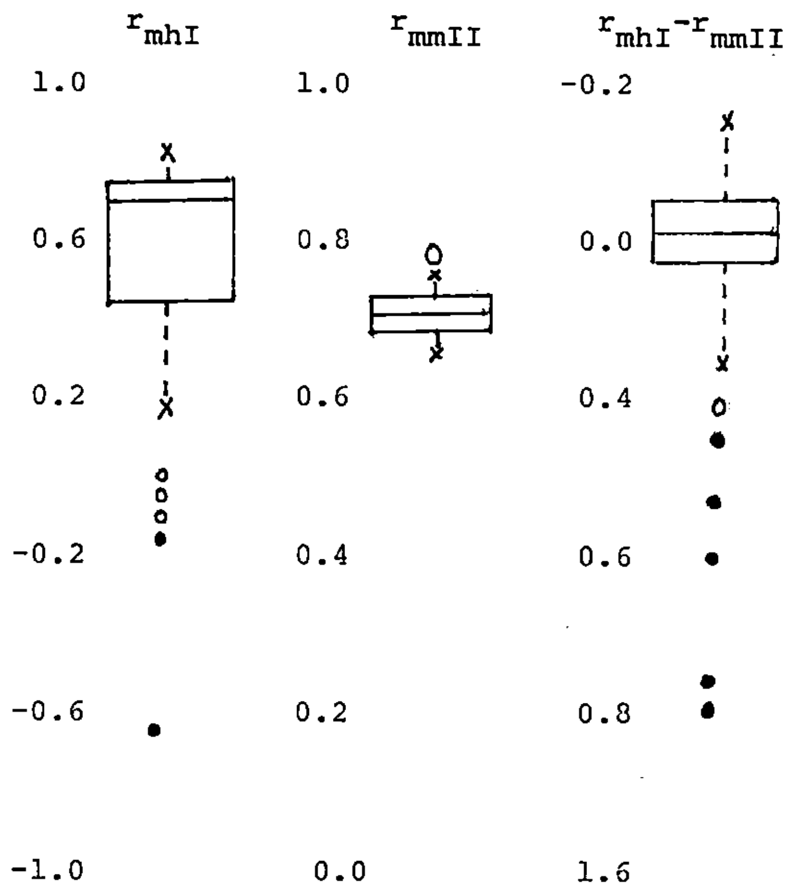
tion accuracy with the "optimal" cue set, $r_{amII} = 0.9127$, and a negative correlation coefficient when selecting the cues, $r_{ahI} = -0.4329$. The box plots indicate that the data contains both moderate and extreme outliers and that the data is not equally distributed around the median.

Table 9, p. 133, is a listing of the prediction accuracies of the participants' models. Nineteen of the participants' models achieved a higher prediction accuracy when the participants selected the cues, Case Set I, than the models using the "optimal" cue set, Case Set II. Figure 6, p. 134, contains the box plots. Again, the mean difference was positive and the median difference was negative, the result of outliers. The data appears not to be symmetrically distributed around the median, but skewed to the right. The distribution of r_{mm} has a slightly lower median, $r_{mmII} = 0.7071$, than the distribution of r_{mhI}, $r_{mhI} = 0.6862$. r_{mmII} is a tighter distribution, with an IQR of 0.0291 versus 0.1885 for r_{mhII}, and has fewer outliers at the lower end of the distribution, five versus two. The two extreme outliers, participants twenty-eight and five, had low prediction accuracies with both models. They are apparently poor decision makers whether the cues are provided or they select their own. The prediction accuracy of the group as a whole, however, appeared to be more consistent when the cues were provided. Other than the outliers, the whole distribution of r_{mm} fits inside

Table 9

PREDICTION ACCURACIES OF THE PARTICIPANTS' MODELS

Participant	r_{mhI}	r_{mmII}	$r_{mhI} - r_{mhII}$
18	0.8110	0.6733	-0.1377
15	0.7304	0.6099	-0.1205
11	0.7659	0.6479	-0.1189
7	0.7443	0.6606	-0.0837
30	0.7578	0.6762	-0.0816
8	0.7354	0.6653	-0.0701
24	0.7921	0.7251	-0.0670
27	0.7598	0.6976	-0.0622
1	0.7450	0.6845	-0.0605
4	0.7618	0.7089	-0.0529
10	0.7285	0.6983	-0.0302
13	0.7217	0.6981	-0.0236
12	0.7150	0.6952	-0.0198
26	0.7071	0.6912	-0.0159
16	0.7009	0.6862	-0.0147
6	0.6912	0.6773	-0.0139
22	0.7178	0.7056	-0.0122
31	0.6944	0.6864	-0.0080
19	0.6754	0.6713	-0.0041
2	0.6921	0.6972	0.0051
23	0.6943	0.7039	0.0096
17	0.7285	0.7412	0.0127
28	0.0449	0.0577	0.0128
20	0.6126	0.6969	0.0843
9	0.4901	0.6587	0.1686
21	0.2969	0.6750	0.3781
25	0.2652	0.6785	0.4133
3	0.1737	0.6879	0.5142
14	0.0966	0.6981	0.6015
5	-0.6248	0.1008	0.7256
29	-0.0964	0.6539	0.7503
Means	0.5590	0.6454	0.0864
Medians	0.7071	0.6862	-0.0139
Difference in the medians	-0.0209		



	r_{mhI}	r_{mmII}	$r_{mhI}-r_{mmII}$
Median	0.7071	0.6862	-0.0139
First Quartile	0.5514	0.6683	-0.0614
Third Quartile	0.7399	0.6974	0.0486
Interquartile Rank	0.1885	0.0291	0.1100
Step	0.2828	0.0437	0.1650
Inner Fence	1.0000	0.7411	-0.2264
	0.2286	0.6246	0.2136
Outer Fence	1.0000	0.7848	-0.3914
	-0.0142	0.5809	0.3786

Figure 6

BOX PLOT: PREDICTION ACCURACY
OF THE PARTICIPANTS' MODELS

the first and third quartiles of the distribution of r_{mh}.

5.1.4 Statistical Results

The statistical results of the tests of the hypotheses are summarized in Table 10, p. 136.[131] There are three potential statistical methodologies available for the testing of the research hypotheses: the sign test, the Wilcoxon signed rank test, and the paired-t test. Each of the methodologies is based on a set of assumptions that must be reasonably satisfied if the procedure is to be applied with confidence. The tests and the critical assumptions required for each procedure were discussed in Chapter Four and are summarized in "Assumptions Required by Tests for Location" in Table 10. The Randle test for symmetry around an unspecified median was used to test for the symmetry of the differences in prediction accuracy; $R_{emI} - R_{ehI}$, $r_{amII} - r_{ahI}$, and $r_{mmII} - r_{mhI}$. The Moses rank-like test for dispersion adapted to paired samples was used to test for the equal dispersion of the two sets of populations; R_{emI} and R_{ehI}, r_{amII} and r_{ahi}, and r_{mmII} and r_{mhI}. Finally, the Kolomogorev-Smirnov procedure was used to test for the normality of the differences in prediction accuracy; $R_{emI} - R_{ehI}$, $r_{amII} - r_{ahI}$, and $r_{mmII} - r_{mhI}$. The p-values

[131] The nonparametric statistics reported in Table 10 were generated using a statistical package developed by Dr. Richard Perie, Professor of Statistics, at Virginia Tech. The appropriateness of the tests and the results of the tests were confirmed by Dr. Perie.

Table 10

SUMMARY OF STATISTICS

Assumptions Required by Tests for Location

Test	Symmetry	Dispersion	Normality
Sign			
Signed rank	x		
Paired-t	x	x	x

P-values for Statistical Tests Performed

	Hypotheses		
Tests for Assumptions	$R_{emI}-R_{ehI}>0$	$r_{amII}-r_{ahI}>0$	$r_{mmII}-r_{mhI}>0$
Symmetry:			
Randle	0.40507	0.00013	0.00092
Dispersion:			
Moses		0.04640	0.02770
Normality:			
Kolmogorev -Smirnov	.15>p>.10	p<.01	p<.01
Tests for Location			
Sign	0.85720	0.89567	0.89567

Assumptions Supported Statistically

Test	Symmetry	Dispersion	Normality
$R_{emI}-R_{ehI}>0$	x		x
$r_{amII}-r_{ahI}>0$			
$r_{mmII}-r_{mhI}>0$			

P-values for Appropriate Tests and Conclusions

Hypothesis	Test	P-value	Conclusion
$R_{emI}-R_{ehI}>0$	Sign	0.81541	No difference
$r_{amII}-r_{ahI}>0$	Sign	0.89567	No difference
$r_{mmII}-r_{mhI}>0$	Sign	0.89567	No difference

of these tests and the critical assumptions that are supported at the 0.05 level of significance are summarized in the second and third sections of Table 10. The results of the tests of the assumptions and the tests of the hypotheses are discussed in detail below.

5.1.4.1 Alternate Hypothesis, Ha: $R_{emII} - R_{ehI} > 0$

It cannot be concluded from the tests of assumptions that the distribution of the differences, $R_{emII} - R_{ehI}$, is not symmetrical, $p = 0.40507$, nor that the distribution of the differences is not normal, $0.15 > p > 0.10$. The Moses test for the equal dispersion of the two populations was not performed because the distribution of R_{emI} is completely defined by its mean, $R_{emII} = 0.7925$. The two distributions, R_{em} and R_{eh}, are, therefore, obviously not equally distributed. (Note the box plots of the data, Figure 4.) The assumptions are met for both the signed rank test and the sign test. Both the signed rank test and the sign test fail to reject the null hypothesis at the 0.05 level of significance, the signed rank test with a p-value of $p = 0.9811$ and the sign test with a p-value of $p = 0.81541$. Both tests conclude that there is no difference between the median prediction accuracy achieved by the participants' environmental models when the participants selected the cues and the prediction accuracy of the environmental models when the cues were provided.

The signed rank test is more powerful than the sign test. Compared to the paired-t test, the asymptotic relative efficiency of the signed rank test is 0.955, and the asymptotic relative efficiency of the sign test is 0.637. Although the sign test is the least powerful of the two tests, it is the test reported for sake of consistency and comparability. As will be seen, the sign test is the appropriate procedure for testing the other hypotheses. The p-value of the sign test is p = 0.81541.

5.1.4.2 Alternate Hypothesis, Ha: $r_{amII} - r_{ahI} > 0$

The results of the tests for assumptions, Table 10, indicate that the sign test is the appropriate test for the above hypothesis. The conclusion of the tests is that the distribution of the differences, $r_{amII} - r_{ahI}$, is not symmetrical and not normally distributed. The Moses test concludes that the distributions, r_{amII} and r_{ahI}, are not equally dispersed. These conclusions are supported by the box plots of the data, Figure 5. The p-value of the sign test is p = 0.89567, and the conclusion is that there is no difference between the groups.

5.1.4.3 Alternate Hypothesis, Ha: $r_{mmII} - r_{mhI} > 0$

Again, the box plots, Figure 6, and the tests for assumptions conclude that the data is not symmetrical, not equally dispersed, and not normally distributed. As a result, the appropriate test is the sign test. The p-value of the sign test is p = 0.89567 and the conclusion is the same, no difference between groups.

To summarize, the tests of hypotheses concluded that there is no difference in the prediction accuracy achieved by the participants when they selected their own cues and the prediction accuracy achieved when the "optimal" cue set was provided. The above results were unexpected. The expectation was that the alternate hypothesis would be accepted; a higher prediction accuracy would be achieved when the participants were provided with a set of cues which were selected using multiple linear regression.

5.2 INFORMATION USE

The results of research using Brunswick's lens model methodology to focus on the accuracy of individual judgments have been consistent in that the prediction accuracy of the environmental multiple linear regression model has invariably been greater than that of both the human decision maker and the multiple linear regression equation representing the decision maker's prediction model. In addition, the studies have in many cases found the prediction accuracy of the decision maker's model to be greater than that of the human decision maker. Theoretically, replacing the human decision maker with his model removes from the prediction process, random error caused by the human decision maker's inconsistent use of his own prediction model.

Replacing the human decision maker with the environmental regression model removes, from the decision process, both random error and systematic bias, caused by the human decision maker's incorrect weighting of the environmental cues. It is therefore hypothesized that prediction accuracy will increase when the human decision maker is replaced by a model because of the elimination of random and/or systematic error from the prediction process. Stated in the alternate form, two hypotheses were tested:

Ha_1: $R_{ehI} > r_{mhI} > r_{ahI}$

Ha_2: $R_{emII} > r_{mmII} > r_{amII}$

5.2.1 Statistical Tests Employed

As explained in Chapter Four, nonparametric procedures are appropriate to test the above hypotheses. The first step in the testing of the above hypotheses is to determine if there is a difference among the participants' prediction accuracies for each of the case sets. A nonparametric randomized complete block ANOVA design was used to test the following hypotheses:

for Case Set I:

Ho: $R_{ehI} = r_{mhI} = r_{ahI}$

Ha: The population medians are not equal.

for Case Set II:

H_o: $R_{emII} = r_{mmII} = r_{amII}$

Ha: The population medians are not equal.

If the null hypotheses are not supported, multiple comparison tests based on Friedman rank sums will be performed to confirm the results of the ordered alternatives tests. The prediction accuracies of the individual participants are in Table 11, p. 142, and the Friedman rank sums are summarized in Table 12, p. 143.

5.2.2 Discussion

Table 11 shows the results of both case sets to be quite similar. The mean prediction accuracies of the environmental models are materially greater than those of both the decision makers' models and the decision makers. There appears to be a small difference between the prediction accuracy of the decision makers and their models. In both Case Set I and Case Set II the mean prediction accuracy of the human decision makers is greater than that of the decision makers' models, the opposite of the hypothesized results.

The Friedman rank sums, Table 12, reflect approximately the same pattern as the means. In Case Set I the prediction accuracy of the environmental model is greater than that of either the participant or the participant's model for twenty-four participants. Seven participants achieved a higher prediction accuracy than their environmental models, and as a result, the total rank sums for the human decision makers were greater than that of the decision makers' models. Seventeen participants achieved a higher prediction accuracy than their models.

Table 11

INDICES OF PREDICTION ACCURACY:
MULTIPLE CORRELATION COEFFICIENTS

Part.	Case Set I			Case Set II		
	R_{ehI}	r_{mhI}	r_{ahI}	R_{emII}	r_{mmII}	r_{amII}
1	0.8243	0.7450	0.8160	0.7787	0.6845	0.6783
2	0.8320	0.6921	0.6212	0.7787	0.6972	0.6650
3	0.6117	0.1737	0.3306	0.7787	0.6879	0.6463
4	0.8313	0.7618	0.6998	0.7787	0.7089	0.7168
5	0.7597	-0.6248	-0.4329	0.7787	0.1008	0.9127
6	0.7469	0.6912	0.7957	0.7787	0.6773	0.6744
7	0.7776	0.7443	0.6904	0.7787	0.6606	0.6014
8	0.7684	0.7354	0.8148	0.7787	0.6653	0.6739
9	0.7302	0.4901	0.5022	0.7787	0.6587	0.6478
10	0.8145	0.7285	0.7961	0.7787	0.6985	0.6976
11	0.8291	0.7659	0.6576	0.7787	0.6470	0.6321
12	0.7855	0.7150	0.7505	0.7787	0.6952	0.6894
13	0.8103	0.7217	0.8049	0.7787	0.6981	0.7087
14	0.8694	0.0966	0.0772	0.7787	0.6981	0.6937
15	0.8288	0.7304	0.7052	0.7787	0.6099	0.5985
16	0.7391	0.7009	0.8079	0.7787	0.6852	0.6911
17	0.8229	0.7285	0.6896	0.7787	0.7412	0.2982
18	0.8577	0.8110	0.7755	0.7787	0.6733	0.6842
19	0.7954	0.6754	0.6458	0.7787	0.6713	0.6576
20	0.7828	0.6126	0.6779	0.7787	0.6969	0.7000
21	0.8238	0.2969	0.1894	0.7787	0.6750	0.6903
22	0.7477	0.7187	0.8225	0.7787	0.7056	0.7064
23	0.7732	0.6943	0.7604	0.7787	0.7039	0.7106
24	0.8445	0.7921	0.7456	0.7787	0.7251	0.5588
25	0.8260	0.2652	0.6610	0.7787	0.7685	0.6761
26	0.7477	0.7071	0.7711	0.7787	0.6912	0.7198
27	0.7730	0.7598	0.8167	0.7787	0.6976	0.6985
28	0.7781	0.0449	-0.1679	0.7787	0.0577	0.3260
29	0.8310	-0.0964	-0.3386	0.7787	0.6539	0.6659
30	0.8657	0.7578	0.7102	0.7787	0.6762	0.6720
31	0.7684	0.6944	0.7734	0.7787	0.6864	0.6853
Means	0.7934	0.5590	0.5797	0.7787	0.6454	0.6574

Differences in means:

$R_{ehI}-r_{mhI}$	0.2344	$R_{emII}-r_{mmII}$	0.1333
$R_{ehI}-r_{ahI}$	0.2137	$R_{emII}-r_{amII}$	0.1213
$r_{mhI}-r_{ahI}$	-0.0207	$r_{mmII}-r_{amII}$	-0.0120

Table 12

INDICES OF PREDICTION ACCURACY:
FRIEDMAN RANK SUMS

Part.	Case Set I			Case Set II		
	R_{ehI}	r_{mhI}	r_{ahI}	R_{emII}	r_{mmII}	r_{amII}
1	3	1	2	3	2	1
2	3	2	1	3	2	1
3	3	1	2	3	2	1
4	3	2	1	3	1	2
5	3	1	2	2	1	3
6	2	1	3	3	2	1
7	3	2	1	3	2	1
8	2	1	3	3	1	2
9	3	1	2	3	2	1
10	3	1	2	3	2	1
11	3	2	1	3	2	1
12	3	1	2	3	2	1
13	3	1	2	3	1	2
14	3	2	1	3	2	1
15	3	2	1	3	2	1
16	2	1	3	3	1	2
17	3	2	1	3	2	1
18	3	2	1	3	1	2
19	3	2	1	3	2	1
20	3	1	2	3	1	2
21	3	2	1	3	1	2
22	2	1	3	3	1	2
23	3	1	2	3	1	2
24	3	2	1	3	2	1
25	3	1	2	3	2	1
26	2	1	3	3	1	2
27	2	1	3	3	1	2
28	3	2	1	3	1	2
29	3	2	1	3	1	2
30	3	2	1	3	2	1
31	2	1	3	3	2	1
Totals	86	45	55	92	48	46

Differences in rank sums:

$R_{ehI}-r_{mhI}$	41	$R_{emII}-r_{mmII}$	44
$R_{ehI}-r_{ahI}$	31	$R_{emII}-ra_{mII}$	46
$rm_{hI}-r_{ahI}$	-10	$rm_{mII}-ra_{mII}$	2

In Case Set II no participant achieved a higher prediction accuracy than the environmental model, and in only one case, participant five, was the environmental model not superior to the participant's model. Fourteen participants had a higher prediction accuracy than their models.

To summarize, in both Case Set I and Case Set II the environmental model was a superior predictor for most of the participants. In general, there appears to be little diference between the prediction accuracy of the participants and their models.

5.2.3 Results of the Statistical Tests

The results of the statistical tests are discussed in two parts. The first section deals with Case Set I test results and the second, with Case Set II test results.

5.2.3.1 Case Set I

The p-value of the Friedman's nonparametric ANOVA test for Case Set I median prediction accuracies is $p < 0.001$. At the 0.001 level of significance the null hypothesis is rejected. There is strong evidence to conclude that at least one median prediction accuracy is different.

Multiple comparison tests using Friedman rank sums confirm the observation that there is little difference in both the means of the prediction accuracies and the Friedman rank sums between the participants and their models. Table 12 summarizes the differences in the rank sums of the alternative comparisons. The test statistic for the multiple comparisons test, with an experimentwise error rate of a = 0.01, is q = 22.94. The difference in the rank sums between the environmental models and the participants' models is forty-one, and the difference in the rank sums between the environmental models and the participants is thirty-one. There is, therefore, strong evidence to conclude that the prediction accuracy of the environmental model is significantly greater than the prediction accuracy of both the participants' models and the participants. The difference in the rank sums of the prediction accuracies between the participants' models and the participants is -10. Therefore, it cannot be concluded that there is a difference in the prediction accuracies of the participants and their models.

To summarize, for twenty-four of the thirty-one participants, the prediction accuracy of the environmental model was greater than the prediction accuracy of both the participants and their models. The multiple comparisons tests support this conclusion at the 0.01 level of significance. An examination of both the means and the rank sums shows little difference in the prediction accuracy of the participants and their models, and it cannot be concluded from the multiple comparisons tests that there is a difference.

5.2.3.2 Case Set II

The p-value of the Friedman's nonparametric ANOVA test for Case Set II median prediction accuracies is $p < 0.001$. At the 0.001 level of significance the null hypothesis is rejected. There is strong evidence to conclude that at least one median prediction accuracy is different.

Again, there is little difference in both the means of the prediction accuracies and the Friedman rank sums between the participants and their models, and, again, multiple comparison tests using Friedman rank sums confirm this observation. The test statistic for the multiple comparisons test, with an experimentwise error rate of $a = 0.01$, is $q = 22.94$. Table 12 shows the difference in the rank sums between the environmental models and the participants' models to be forty-four, and the difference between the environmental models and the participants to be forty-six. There is strong evidence to conclude that the prediction accuracy of the environmental model is significantly greater than the prediction accuracy of both the participants' models and the participants. The difference in the prediction accuracies between the participants' models and the participants is two; it cannot be concluded that there is a difference in the prediction accuracies of the participants and their models.

The results of the statistical tests and the conclusions are the same as for Case Set I. The prediction accuracy of the environmental model was greater than the prediction accuracy of both the participants and their models for all but one participant. The multiple comparisons tests support this conclusion at the 0.01 level of significance. Both the means and the rank sums show little difference in the prediction accuracy of the participants and their models, and it cannot be concluded from the multiple comparisons tests that there is a difference. Table 13, p.148 summarize the statistical tests and the conclusions.

5.3 INFORMATION OVERLOAD

As discussed in Chapter three, each participant initially performed the experiment using two cues to generate an estimate of each firm's prediction accuracy. The participant then performed three additional iterations of the experiment receiving two additional cues for each new iteration until a total of eight cues was used in the final iteration and four predictions of the environmental event were generated.

5.3.1 Definition of Terms

Two terms, prediction accuracy and confidence, are essential to the discussion of information overload. Their meanings for the purposes of this study are explained below.

5.3.1.1 Prediction Accuracy

As in Case Set I and Case Set II the prediction accuracy of the human decision maker is defined as the Pearson Product Moment Correlation Coefficient between the environ-

Table 13

INFORMATION USE: SUMMARY OF STATISTICS

	Case Set I	Case Set II
Treatments		
1	R_{ehI}	R_{emII}
2	r_{mhI}	r_{mmII}
3	r_{ahI}	r_{amII}
Descriptive statistics		
Means		
R_e	0.7934	0.7787
r_m	0.5590	0.6454
r_a	0.5797	0.6574
Differences in means		
$R_e - r_m$	0.2344	0.1333
$R_e - r_a$	0.2137	0.1213
$r_m - r_a$	-0.0207	-0.0120
Friedman rank sums		
R_e	86	92
r_m	45	48
r_a	55	46
Differences in rank sums		
$R_e - r_m$	41	44
$R_e - r_a$	31	46
$r_m - r_a$	-10	2

Table 13 (Continued)

Statistical Results	Case Set I	Case Set II
ANOVA test		
Hypotheses		
Ha_1: $R_{ehI} > r_{mhI} > r_{ahI}$		
Ha_2: $R_{emII} > r_{mmII} > r_{amII}$		
P-values	p 0.001	p 0.001
Decisions	Accept Ha_1	Accept Ha_2
Multiple comparisons tests		
Experimentwise error rate	a = .01	a = .01
Test statistic	q = 22.94	q = 22.94
Conclusions		
R_e and r_m are different		
R_e and r_a are different		
r_m and r_a are not different		

mental event, each firm's ROA, and the participants' predictions of the environmental events. Four correlations were therefore generated for each participant; r_{ah2}, r_{ah4}, r_{ah6}, and r_{ah8}. r_{ah2} is the prediction accuracy the participant achieved using two cues. r_{ah4}, r_{ah6}, and r_{ah8} are the prediction accuracies achieved by the participants using four, six and eight cues respectively.

5.3.1.2 Confidence

The participants' confidence in each of their predictions was captured using a Likert scale. Subsequently, the participant's average confidence level was computed for each of the four iterations of the experiment. CL_2 is the average confidence level the participant reported when using two cues to predict ROA. CL_4, CL_6, and CL_8 were the average confidence levels reported by the participants using four, six, and eight cues respectively.

5.3.2 Results of the Statistical Tests

Two key factors, prediction accuracy and confidence level, relate to information overload. These factors were tested statistically.

5.3.2.1 Prediction Accuracy

It was hypothesized in Chapter Three that:

H_{a6}: Prediction accuracy (r_{ah}) will be maximized after the decision maker's examination of a small number of cues and then decline as additional cues are examined.

The first step in the testing of this hypothesis is to determine if there is a difference among the prediction accuracies the participants achieved using two, four, six, and eight cues. A parametric randomized complete block ANOVA model was used to test:

Ho: $r_{ah2} = r_{ah4} = r_{ah6} = r_{ah8}$

Ha: The population means are not equal.

The p-value for the ANOVA procedure was $p = .0528$. It was therefore concluded, at the .0528 level of significance, that at least one mean prediction accuracy was different.

The ANOVA procedure indicates that there are differences in the prediction accuracies achieved by the participants using different numbers of cues. Multiple comparison procedures are therefore appropriate to determine the specific differences. The expectation is that prediction accuracy will first increase, then decrease as the decision maker receives additional cues. The specific differences of interest are the differences between r_{ah2} and r_{ah4}, r_{ah4} and r_{ah6}, and r_{ah6} and r_{ah8}.

Several multiple comparison procedures are available, and these are discussed by Kleinbaum and Kupper and compared to the Least Significant Difference method.

> A disadvantage of the LSD method, however, is that the true overall significance level may be so much less than the maximum value a that none of the individual tests is very likely to be rejected (i.e., the overall power of the method is low). Consequently, several more powerful procedures have been devised which can be used

to provide an overall significance level . . .[133]

Kleinbaum and Kupper also state that either Tukey or Scheffe tests may provide more precise results, and they delineate the criteria for the proper use of the Tukey methodology.

> Tukey's method is applicable when:
> 1. The sizes of the samples selected from each population are equal. . .
> 2. Pairwise comparisons of the means are of primary interest. . .[134]

Tukey's procedure is an appropriate method for this study because the sample size for each treatment is the same, thirty-one, and pairwise comparisons, the differences between r_{ah2} and r_{ah4}, r_{ah4} and r_{ah6}, and r_{ah6} and r_{ah8}, are of primary interest. An advantage of Tukey's multiple comparisons procedure is that the experimentwise significant level is known. Neter, Wasserman, and Kutner state, "The Tukey method is exact when all sample sizes are equal. . ."[135]

The results of the multiple comparisons procedures are presented in Table 14, p. 153. As can be seen in the graphical representation, the prediction accuracies follow the hypothesized pattern. Prediction accuracy increases

---

[133]D. G. Kleinbaum and L. L. Kupper, Applied Regression Analysis and Other Multivariate Methods, (North Scituate: Duxbury Press, 1981), p. 265.

[134]Ibid., p. 268.

[135]John Neter, William Wasserman, and Michael H. Kutner, Applied Linear Statistical Models (Homewood, Illinois: Richard D. Irwin, Inc., 1985). p. 574.

Table 14

PREDICTION ACCURACY:
TUKEY'S MULTIPLE-COMPARISON PROCEDURE

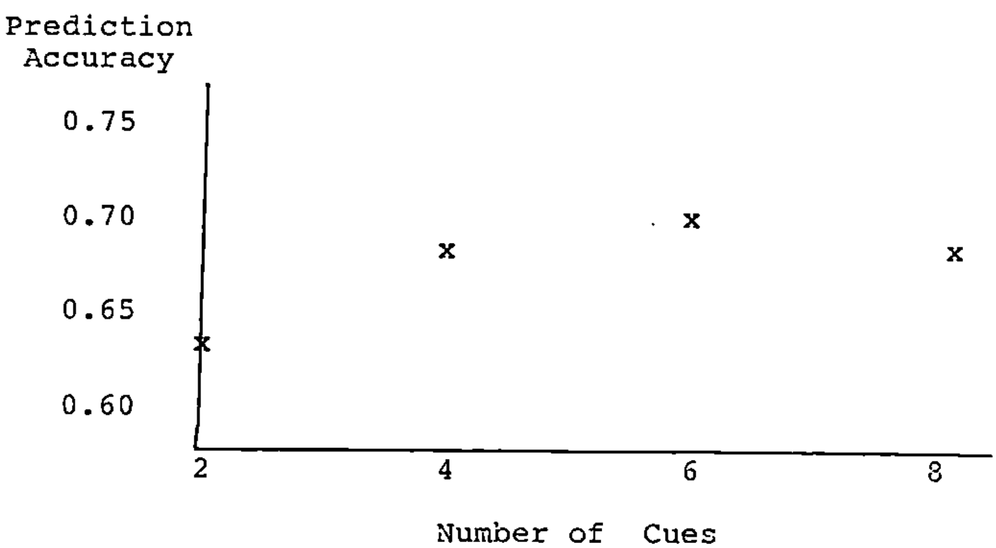
Alpha = 0.05
Minimum significant difference is 0.07366

Number of Cues	Mean Prediction Accuracy
2	0.63567
8	0.68935
4	0.69110
6	0.71138

Results

2 8 4 6

Graphical Representation of the Results



as the number of cues increases from two to six and then decreases when the number of cues increases to eight. Hypothesis Ha_7, information overload, however, is not supported statistically at the .05 level of significance. Statistically, there is no difference between r_{ah2}, r_{ah8}, and r_{ah6}; and r_{ah8}, r_{ah4} and r_{ah6}. A statistical difference does exist between r_{ah2} and r_{ah6}. The decrease in mean prediction accuracy between r_{ah6} and r_{ah8}, required for the acceptance of Ha_6, was not supported statistically.

5.3.2.2 Confidence Level

The hypotheses relating to the participants' expected levels of confidence as they received more information is:

Ha_7: The decision maker's confidence in his prediction accuracy will increase with the use of additional cues.

The procedure for testing this hypothesis is the same as that used to test Ha_6. If the ANOVA F-test used for simultaneously comparing the several population means is significant, Tukey's multiple comparison procedure will be used to test for specific differences. The discussion of the appropriateness of the parametric randomized complete block ANOVA methodology and Tukey's procedure is equally applicable to the test of Ha_8 and will not be repeated.

The specific hypotheses tested using the parametric randomized complete block ANOVA model are:

Ho: $CL_2 = CL_4 = CL_6 = CL_8$

Ha: The population means are not equal.

The p-value for the ANOVA procedure was $p = 0.0006$. It was therefore concluded, at the 0.0006 level of significance, that at least one mean confidence level was different. The expectation was that the decision makers' confidence in their decisions would increase as they received additional cues to be used as information upon which to base their predictions. The specific differences of primary concern, therefore, are CL_2 and CL_4, CL_4 and CL_6, and CL_6 and CL_8.

The results of the multiple comparisons are presented in Table 15, p. 158. As can be seen in the graphical representation of the results, the mean confidence levels follow the hypothesized pattern. The participants' confidence in their decisions continues to increase as they receive additional information. Hypotheses Ha_7, however, is not supported statistically. At the .05 level of significance there is no difference between CL_4, CL_6 and CL_8. A statistical differnce exists between CL_2 and CL_4, CL_2 and CL_6, and CL_2 and CL_8. The participants' confidence levels significantly increase upon receiving four cues and then remain constant.

5.4 POST-TEST QUESTIONNAIRES

A separate post-test questionnaire was employed for

Table 15

CONFIDENCE LEVEL:
TUKEY'S MULTIPLE-COMPARISON PROCEDURE

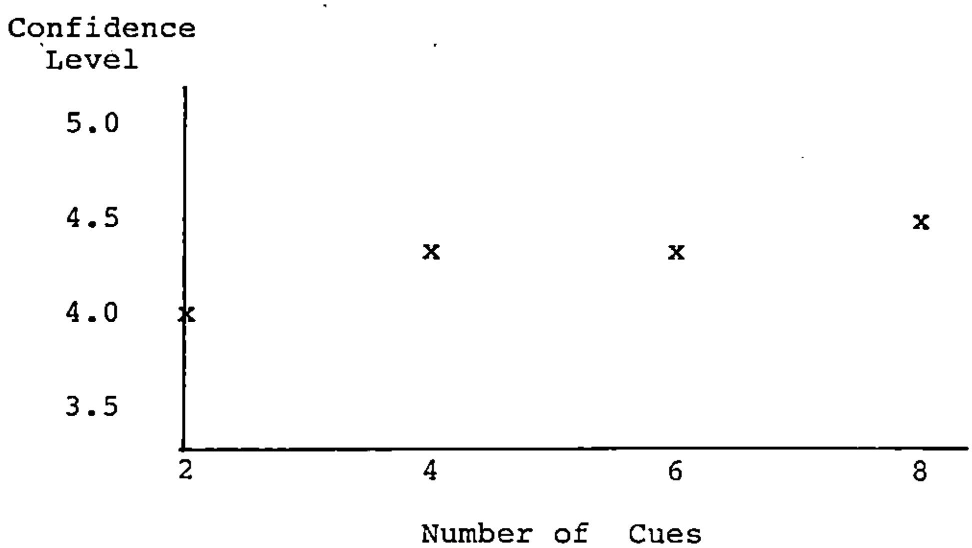
Alpha = 0.05
Minimum significant difference is 0.3839

Number of Cues	Mean Confidence Level
2	3.9667
4	4.4677
6	4.4677
8	4.5312

Results

2 4 6 8

Graphical Representation of the Results



each experiment. The questionnaires were used to capture demographic data about the participants, to gain insight into their attitudes towards the study, and to determine the level of their motivation to perform well on the experimental tasks.

5.4.1 Purpose of the Questionnaires

The purpose of the questionnaires was to determine the factors that were associated with the participants' performance of the experimental tasks. Specifically, the information gathered by the questionnaires was used in an attempt to answer four questions. The first three questions relate to both experiments, and the fourth is applicable to the information overload experiment only. The questions are:

1. Is there a relationship between age, academic background, and other demographic and background factors and the participant's prediction accuracy?
2. Do attitudinal factors affect prediction accuracy?
3. Do the participants who reported that they were stimulated to perform by the possibility of financial reward achieve a higher prediction accuracy than those who reported that they were not stimulated by the possibility of financial reward?
4. Do the participants have an intuitive understanding of the information overload phenomenon? Specifically, do the participants recognize the point at which

they maximize prediction accuracy?

The questionnaires were also examined to determine what other factors might impact on prediction accuracy. The development and use of heuristics by the participants and their level of understanding of the experimental task were two factors identified.

Because the central purpose of the post-test questionnaire is to provide background information for the study and to identify extraneous factors that might impact on the results of the experiment, an in depth statistical analysis of the questionnaire was not undertaken. The attitudinal factors captured using a Likert scale were the only questionnaire items analyzed statistically.

5.4.2 Cue Selection/Cue Use Questionnaire

The cue selection/cue use questionnaire was employed to determine the impact of various factors on the prediction accuracy achieved by the participants in the experiment. Among those factors examined were the participants' backgrounds, their attitudes toward the experiment, and their motivation for achieving a high level of prediction accuracy. A sample of the questionnaire is in Appendix A.

5.4.2.1 Affects of Background on Prediction Accuracy

Five factors relevant to the participant's background were examined to determine their affect on the prediction accuracy achieved by the participant. The five factors were educational status (Masters of Accountancy or MBA), age, study financial statement analysis, work experience in accounting or finance, and expected field of employment. Intuitively, it was expected that participants enrolled in the Masters of Accountancy program, older participants, participants who had studied financial statement analysis, those with work experience in professional accounting or finance, and those who expected to enter the fields of accounting or finance, would have an advantage in predicting a firm's future ROA. If they did have an advantage, a disproportionate number of these participants would have achieved a prediction accuracy at or above the median in both Case Set I and Case Set II. (Prediction accuracy is defined here as the prediction accuracy of the participants, r_{ahI} for Case Set I and r_{amII} for Case Set II.) The results, summarized in Table 16, p. 160, are as follows:

1. Ten of the thirty-one participants were enrolled in the Masters of Accountancy program. Six of these exceeded the median prediction accuracy in both Case Set I and Case Set II.
2. Nine participants were over thirty years old. Three exceeded the median prediction accuracy of Case Set I, and four exceeded the median prediction accuracy of Case Set II.
3. Sixteen participants answered yes to the question, "Have you ever had a course in financial statement analysis?" Five of the sixteen had a prediction

Table 16

CUE SELECTION/CUE USE
POST-TEST QUESTIONNAIRE:
ANALYSIS OF BACKGROUND INFORMATION

Quest.	Description	Number of Respondents At or Above Median Prediction Accuracy $\underline{r}_{ah}$	$\underline{r}_{am}$
*	Ten participants were enrolled in the Masters of Accountancy program.	6	6
1	Nine participants were over the age of 30.	3	4
2	Sixteen participants reported completing a course in financial statement analysis.	5	8
3	Six participants reported prior job experience in accounting or finance.	3	3
4	Eighteen participants expected to obtain a position in either accounting or finance upon graduation.	10	10
6	Twenty-five participants found the study educational.	10	13
7	Twelve participants indicated that the potential for reward stimulated their performance.	6	5
8	Twenty-one participants reported developing a strategy prior to the selection of the cues for Case Set I.	13	12
9	Of the ten participants who did not develop a strategy prior to the selection of the cues for Case Set I, six developed a strategy after selecting the cues.	2	3

Table 16 (Continued)

10	Thirty participants responded to question 10. Twenty-nine indicated that they would be willing to participate in such an exercise again.	15	15
12	Thirty participants responded to question 12, twenty-nine of whom stated that they would like a summary of their prediction accuracies.	15	15

accuracy above the median of Case Set I, and eight had a prediction accuracy above the median for Case Set II.

4. Six participants reported job experience in professional accounting or finance. One participant reported nine years' work experience; one, five years' experience; and four, two years' experience or less. The prediction accuracy of three of these exceeded the median prediction accuracy of both the case sets. Participant number two reported nine years of work experience and ranked twenty-fourth in Case Set I and twenty-second in Case Set II, well below the median prediction accuracy in both case sets. Participant number twenty-seven reported over five years of work experience and ranked second in Case Set I and eighth in Case Set II.
5. Ten of the eighteen participants planning to obtain a position in either accounting or finance achieved a prediction accuracy above the median in both case sets.

There appears to be no relationship between the above factors and prediction accuracy.

5.4.2.2 Attitudinal Factors

Question five on the Cue Selection/Cue Use questionnaire used a seven point Likert scale to capture three of the participants' attitudes towards the study; boring or interesting, fatiguing or stimulating and unrealistic or realistic. The Pearson Product Moment Correlation Coefficient was used to determine if there was a relationship between these attitudinal factors and prediction accuracy. Table 17, p. 164, shows no significant correlation between the attitudes and the participants' performance.

Question six asked the participants if they found the experiment to be educational or not educational. Twenty-five participants stated that the experiment was an educational experience. Ten of these twenty-five participants achieved a prediction accuracy above the median in Case Set I, and thirteen achieved a prediction accuracy above the median in Case Set II.

Question ten asked, "Would you volunteer to participate in such an exercise again?" Question twelve asked, "Would you like a summary of your prediction accuracy?" Twenty-nine participants answered yes to both questions, one participant answered no to both questions, and one did not answer the question. The affirmative answers to these questions indicate a high level of interest in the experiment.

Questions five, six, ten, and twelve attempted to capture the participants' attitudes towards the study. The analysis of the participants' responses to these questions indicate that there is no relationship between attitudinal factors and prediction accuracy.

5.4.2.3 Motivation

Table 17

CUE SELECTION/CUE USE
POST-TEST QUESTIONNAIRE:
CORRELATION OF ATTITUDE CHARACTERISTICS
AND PREDICTION ACCURACY

Characteristic	Correlation	
	Case Set I	Case Set II
Boring/Interesting	-0.1170	-0.2730
Fatiguing/Stimulating	-0.0660	-0.2110
Unrealistic/Realistic	-0.1360	-0.2180

Question seven attempted to determine if the potential for reward stimulated a higher level of performance. Twelve participants indicated that their performance was stimulated by the reward structure. Six of the twelve participants had a prediction accuracy above the median in Case Set I, and five, above the median in Case Set II. The participants who reported that the potential for a reward stimulated them to perform to the best of their abilities did not appear to achieve consistently higher prediction accuracies than those who did not.

Motivation to perform may have been innate in many of the participants. Participant number eight answered question seven no and stated, "[My] Motivation: [was based on a] sense of responsibility to fellow graduate students." Participant number fourteen answered no, "but I did feel like I was taking a test and wanted to give the correct answer."

5.4.2.4 Other Factors

Questions eight and nine asked the participants if they had devised a prediction strategy and if so, to describe the strategy. The participants' responses are summarized in Table 18, p. 166. From the participants' descriptions the one clear strategy used was anchoring and adjustment. Judgmentally, twelve participants were identified as using this heuristic.

The most common anchoring point reported is the current

Table 18

CUE SELECTION/CUE USE QUESTIONNAIRE:
ANSWERS TO QUESTIONS EIGHT AND NINE
THE PARTICIPANTS' REPORTED PREDICTION STRATEGIES

Part.	Question Number	Description of the Strategy
1	9	Beg. point - ROA then adj. from there based on other ratios
2	8	(Net income/total assets). Operating income to total assets - Sales growth (debt/assets), Industry return on assets.
3	8	I multiplied NI/S times S/TA and compared to ROA asset growth and income growth
4	8	Picked out cues that seemed relevant to est. ROA -
5	NO RESPONSE	
6	8	Assume similar % growth
7	8	I looked at WC (2) for firm "health";(14) for a bench mark, (23) for "health" (20) and (26) to determine a trend for the firm in relation to the industry; and (18) for determining what would happen to (14) in the future.
8	8	Cues 4 & 5 to indicate business specific problems. 18, 19, 20 to indicate trends in ROA. 26 to indicate industry growth or decline.
9	NO RESPONSE	
10	8	Tried to find ratios that would reveal ROA.
11	9	Would have used cue 5 & cue 10 to get a rough estimate, if I had had cue 5.
12	8	Examine 1984 ROA, Examine change in ROA over 1983-84, Examine change in sales and changes in industry.
13	9	Determining a growth rate for the 19x3-19x4

Table 18 (Continued)

		years, adjusting the 19x4 ROA (for this), then readjusting based on other data.
14	9	Looked for market share (Hi or Low) Based predicted ROA on combination of: net income/sales revenue; sales/oper. assets; oper. income/total assets. Predicted higher than industry ROA if all above rel. high & high mkt share. Predicted lower than industry ROA if all above rel. low and/or low mkt share.
15	8	Concentrated on Oper Inc/Oper Asset (ratio 14) as basis then roughly adjusted for 19x5 with Op Inc Growth (18) & Op. Asset Growth (19).
16	8	Really little maybe. (1) am't of debt (2) return to stockholders might indicate management's forecast for success the next year (3) ROA of previous year
17	9	Based prediction on operating income/total assets ratio. Modified by trends in income growth and operating asset growth.
18	8	Used trends as best as possible from both industry and firm & relate them to my answer. Also, if trends were opposite in direction I put less faith in my answer.
19	8	Some of the cues were selected for Expected Return and some were selected for variability of the return. (The more variable, the less confidence in E(R).)
20	8	Pick ratio with Op Income or Op Assets. Pick growth stats w/ these factors. Pick present ROA.
21	8	Part of my strategy was based on looking at sales growth of previous years.
22	8	To see what the current rate of return, then based on the operating income growth rate and operating asset growth rate predict the ROA for 19x5. I also used operating

Table 18 (Continued)

		income/sales rev. and sales/operating assets to aid me in my prediction.
23	8	Selected one that seemed to be a good predictor from each category, then eliminated one from the "Other Ratios" and selected another from "Trends."
24	9	I compared the industry ROA to the firm's ROA from the previous year & predicted the current ROA based on the closeness of the two. (I soon realized that it would have helped immensely if I had chosen some trend ratios.
25	NO RESPONSE	
26	8	Look at ROA for present year, then look at history of operating income growth & capital spending growth to determine if the present year would be higher or lower.
27	8	Basically to determine how the other ratios that I picked looked in relation to what I thought would predict ROA best.
28	8	Select cues such that the ratios might provide a factor(s) to be used to help ascertain an expected (predicted) ROA.
29	8	Used: Oper income/Oper assets. ROA for the industry.
30	8	Check asset growth, income growth - factor in market share and market growth and sales.
31	8	Attempt to identify current ROA, how impt. sales were to ROI & if sales were rising or falling compared to industry sales.

period's ROA. Participant seventeen described one personal strategy as follows: "[I] Based [my] prediction on [the] operating income/total assets ratio [and] modified [it] by trends in income growth and operation asset growth." A second possible anchoring point is cue number twenty-nine, industry ROA, which was selected by sixteen of the thirty-one participants. No participant, however, explicitly reported using industry ROA as an anchor. Of the twelve participants who apparently used an anchoring and adjustment heuristic, seven achieved a prediction accuracy above the median in Case Set I. There is no clear indication that the participants who used an anchoring and adjustment heuristic had a superior prediction accuracy.

Question eleven asked, "If you could give one piece of advice to the experimenter, what would it be?" The participants' responses to this question were scrutinized for any other factors that might have affected experimental performance. The participants' answers to question eleven are recorded in Table 19, p. 170. There is no clear pattern to the responses to this question. Several participants would have liked more information; however, there is no consistency to the type of information requested. Participant seventeen suggested, "Indicate the type of industry the firms are in." Participant twenty-five stated, "Ratio units should be given in screen menu (times, days, percentages, etc.)"

Table 19

CUE SELECTION/CUE USE QUESTIONNAIRE:
ANSWERS TO QUESTION ELEVEN
PARTICIPANTS' ADVICE FOR THE EXPERIMENTER

Part.	Comment
1	None.
2	NO RESPONSE
3	The ability to ask the experimenter questions at any time really eliminates the need for any changes in the procedure.
4	NO RESPONSE
5	Would give participants more verbal direction on purpose and <u>need</u> for strategic planning @ beginning.
6	Indicate the difficulty of forecasting; I felt my background was inadequate for the task & consequently I felt "dumb."
7	Please explain goals before we begin experiment - also - try writing a programmed-learning "book" or "set" with this material - it is fun to do and painless -
8	NO RESPONSE
9	Have the volunteers spend 5 minutes or more reading the instructions before entering the terminal area.
10	Schedule me in the morning.
11	Give further explanation of the ratios (if it wouldn't defeat the purpose).
12	NO RESPONSE
13	You don't have to pay $10 - I would have come in for free.
14	Give a better idea of what cues will be used before the first set of cases. Would give a better chance to choose the most appropriate cues in the first place.

Table 19 (Continued)

15	Choose his initial set of 6 ratios for the 1^{st} set of cases carefully. many of the 30 ratios appear irrelevant in prediction of ROA.
16	Choose as uniform a background as possible for your subjects.
17	Indicate the type of industry the firms are in.
18	To recommend the use of a calculator during the study.
19	I would be interested to see the results of a 3^{rd} set of cues which include the user's initial 6 cues with the other 6 cues.
20	(1) Give written formula of each ratio. (2) Use same wording for all ratios whereever they appear.
21	Instead of 30 firms, 20 firms' data would be appropriate, because one gets tired looking at the numbers after a certain time.
22	No advice, he did an excellent experiment and it was very interesting.
23	Provide a display screen at the bottom of display to perform manipulations with ratios; cut down on number of questions - gets boring after 20.
24	Make sure we understand exactly what we're supposed to do before we begin. It took me a few entries to understand what to do.
25	Ratio units should be given in screen menu (times, days, percentages, etc.)
26	NO RESPONSE
27	Use more than one set of ratios, perhaps the past 5 years so that one could get a better feel for the trend of the companies.
28	No criticisms or comments seem necessary.
29	NO RESPONSE
30	A short rundown of some aspects of each firm - type

Table 19 (Continued)

	of products, etc. would lead to better use of ratios. General information about the firm could alter use of ratios in predicting future ROA.
31	Had trouble with the confidence scale - maybe it could be constructed with better anchors - what is "average" confidence? (#4)

Most participants appeared to understand the experimental task. Only two, number twenty-four and number thirty-one stated explicitly that they were confused. Participant twenty-four wrote, "Make sure we understand exactly what we're supposed to do before we begin. It took me a few entries to understand what to do." Participant number thirty-one stated, "[I] had trouble with the confidence scale - maybe it could be constructed with better anchors - what is 'average' confidence? (#4)" Both participants performed above the median prediction accuracy on Case Set I.

5.4.3 Information Overload Questionnaire

The post-test questionnaire for the information overload portion of the study followed the same format and captured the same data as the Cue Selection/Cue Use Questionnaire. The examination of Tables 20, 21, 22, and 23, beginning on p. 174, indicates that the conclusions drawn from the Cue Selection/Cue Use Questionnaire are equally applicable to the Information Overload Questionnaire. There appears to be no relationship between background, attitudinal and motivational factors and prediction accuracy.

In addition to the above factors, the Information Overload Questionnaire attempted to determine if the participants had an intuitive understanding of the information overload phenomenon. Table 24, p. 182, summarizes the

Table 20

INFORMATION OVERLOAD
POST-TEST QUESTIONNAIRE:
ANALYSIS OF BACKGROUND INFORMATION

Quest.	Description	Number of Respondents At or Above Median Prediction Accuracy
*	Ten participants were enrolled in the Masters of Accountancy program.	4
1	Seven participants were over the age of 30.	2
2	Nineteen participants reported completing a course in financial statement analysis.	9
3	Three participants reported prior job experience in accounting or finance.	2
4	Eighteen participants expected to obtain a position in either accounting or finance upon graduation.	9
6	Twenty-two participants found the study educational.	10
7	Ten participants indicated that the potential for reward stimulated their performance.	5
9	Sixteen participants reported developing a strategy prior to the selection of the first two cues.	9
10	Of the fifteen participants who did not develop a strategy prior to the selection of the first two cues, thirteen developed a strategy after selecting the first two cues.	6
12	Twenty-eight participants responded to question 12. Twenty-five	

Table 20 (Continued)

	indicated that they would be willing to participate in such an exercise again.	12
14	Twenty-eight participants responded to question 14, nineteen of whom stated that they would like a summary of their prediction accuracy.	11

Table 21

INFORMATION OVERLOAD
POST-TEST QUESTIONNAIRE:
CORRELATION OF ATTITUDE CHARACTERISTICS
AND PREDICTION ACCURACY

Characteristic	Correlation
Boring/Interesting	-0.1950
Fatiguing/Stimulating	-0.2990
Unrealistic/Realistic	-0.4320

Table 22

INFORMATION OVERLOAD QUESTIONNAIRE:
ANSWERS TO QUESTIONS EIGHT AND NINE
THE PARTICIPANTS' REPORTED PREDICTION STRATEGIES

Part.	Question Number	Description of the Strategy
1	10	I used the firm's 1984 number as a basis and then checked sales growth and capital spending growth. As more ratios were selected, I checked how these factors might influence the number previously obtained.
2		No reported strategy.
3	9	Tried to look for relation in cues and priority ordering.
4	10	For the first two cues I used (#3 & 5), I screwed up. However, the next two (#6 & 7) proved to be exactly what I needed. I would compare the increase (decrease) in operating income w/ the increase (decrease) in capital spending, and use this comparison to judge if I thought the ROA would increase or decrease.
5	9	I selected the two believed to be most important in the prediction.
6	9	Figure firm ROA & potential NOI growth as best measures.
7		No reported strategy.
8	9	Look at 19x4 ROA & predict 19x5 ROA based on change in income and capital spending.
9	9	My strategy at that point was to rank the cues in order of importance.
10	9	Find a point estimate of ROA (84 ROA) and see what growth rates were for the parts of the ratio.
11	10	I simply tried to use the figures which I thought best represented the potential for performance estimation for the firm.

Table 22 (Continued)

		I think I may be biased towards the importance of capital investment, & so chose it first.
12	9	According to the ratio OpI/OpA -- see increase in Inc and effects on ratio.
13	10	(1) Look at Cue # 3 & Cue # 5 to see how related. (2) Check Cue # 6 & Cue # 7 -- Check Cue # 4.
14	10	General trend of what was happening for a firm -- would have helped to have previous ROA we made shown -- would have made it more accurate seeming & less fatiguing.
15	10	Started to estimate by ranking the cues.
16	10	Tried to compare capital spending growth with sales growth in relationship to operating income growth for years 2 through 4. Compared Industry Sales Growth with that of the firm.
17	10	Tried to get cues that would best estimate the components of ROA.
18	9	Attempted to solve for ROA using various ratios, w/ a Dupont Anal in mind. Lost track w/ fatigue.
19	9	Rely heavily on x4 ROA figure, avoid high confidence levels.
20	9	Compare initially to industry, then get info. on sales growth for both. Also, don't think the ratios will change radically.
21	10	Tried to compare trends (sales, ROA) of industry v. trends (sales, ROA) of firm, decide whether investing in assets or letting assets "expire" or if trying to boost ROA by <u>not</u> investing.
22	9	Pick the 2 cues that compliment each other, not necessarily the 2 best out of the 8.
23	9	1^{st} cues 3 & 6, then 4 & 5, then 7 & 8,

Table 22 (Continued)

		then 1 & 2.
24	9	Look at past year's ROA and recent sales growth and expect a trend to continue.
25	10	Prioritizing the remaining cues.
26	10	I concentrated on firm ROA - 19x4, and then considered the operating income growth/capital spending growth.
27	9	Chose 19x2-x4 operating income, ROA for 19x4, decided the relation would give some hint. After developing initial strategy, I completed about half of the first part (two cues), then decided that additional info would help, so chose industry ROA, sales growth for the next step, capital spending, industry sales for the next one.
28	10	100% increase in capital spending -- 25% decrease in ROA if income remains stable. Later discarded the strategy.
29	9	Look at 1984 ROA and adjust up or down.
30	10	(1) Balance increase/decrease in net income for 19x1--19x4 against increase/decrease in capital spending growth. (2) Figure out which of the 2 exceeded the other in growth. (3) Assume the trend will continue. (4) Adjust the ROI for the trend.
31	9	Used the elements of ROA (cues 3 & 5), and history of ROA to go by.

Table 23

INFORMATION OVERLOAD QUESTIONNAIRE:
ANSWERS TO QUESTION THIRTEEN
PARTICIPANTS' ADVICE FOR THE EXPERIMENTER

Part.	Comment
1	--Suggest that everyone choose the 19x4 cue in the set or predictions will be way off- --Tell those doing study that cues will be added to ones chosen in previous step--I didn't figure that out.
2	Some of the case data was unrealistic; i.e. firm data was significantly different from industry data in almost all cases.
3	Better instructions for use of cues.
4	Maybe cut down on the number of companies.
5	Make the pages so that I can go back to a previous page because I believe I skipped the confidence level in one of the cases. Too bad I couldn't go back and fill it out.
6	Base answers primarily on past firm ROA if you are the subject. Experimenter--try to pick a testing room that offers a few more comforts.
7	Get finance people.
8	Have the program ask you what order you want the cues listed in.
9	I don't have any advice--sorry! I enjoyed the experiment & am curious about how I predicted.
10	NO RESPONSE
11	NO RESPONSE
12	To have a longer list of information to select about the firms with 8 runs, but a choice of 2 new items each time.
13	Tell people to read directions through first before studying the cue list.

Table 23 (Continued)

14	Give previous prediction ROA on next cues.
15	Perhaps conduct the experiment at an earlier time of the quarter.
16	NO RESPONSE
17	NO RESPONSE
18	NO RESPONSE
19	I think everyone participating would have been interested in finding out more about the study and your dissertation.
20	Possibly give the nature of the industry to aid prediction--more realistic.
21	More comparable data; e.g. sales growth for x4 w/ industry sales growth for x4 rather than industry x4 v. firm growth x2-x4.
22	Make it clear that the cues looked at will total 8 for the last prediction. That is that the cues will be cumulative.
23	Try to arrange participants a little more in advance, especially near finals.
24	In most industries an estimate of sales or general trend can be obtained--I would look at these also.
25	Have no advice--study seemed to be under control.
26	Perhaps we should take the test without giving our names.
27	I think I didn't quite understand experiment format until I got into the program. Would have been helpful to explain exactly what was going on--but maybe I missed it. I was late.
28	NO RESPONSE
29	Don't schedule 3 days before comps.
30	Have it at the beginning of the quarter when everyone has more time.
31	Too much info to put together beyond 4 cues.

Table 24

INFORMATION OVERLOAD QUESTIONNAIRE:
ANSWERS TO QUESTION ELEVEN
PARTICIPANTS' INTUITIVE UNDERSTANDING OF
INFORMATION OVERLOAD

Number of Cues Required to Predict ROA	Number of Participants Who Indicated Prediction Accuracy Would Not Improve
2	-
4	13
6	7
8	1
More than 8	1

participants' responses to question eleven, "Do you believe that your prediction accuracy improved, or would improve, as a result of continuing the experiment beyond the number of cues indicated above?" Twenty-two of the twenty-eight participants responding to the question replied in the negative. Their negative responses indicate that these participants did have an intuitive understanding of the information overload phenomenon.

5.4 SUMMARY

There were two major objectives of this study, both based on Brunswick's lens model methodology. The first major objective was to determine if forecasts employing environmental cues selected by a mathematical model are more accurate than forecasts based on cues selected by human decision makers. It was expected that the mathematical selection of cues would yield more accurate forecasts because the use of multiple linear regression to select the cues would generate the "optimal" set of cues and weight them so as to maximize prediction accuracy.

The results obtained, however, were unexpected. The median prediction accuracies of the environmental models, the participants, and the participants' models were greater when the cues were selected by the participants. This result, however, was not supported statistically. The tests of the hypotheses concluded that there was no difference in the prediction accuracies achieved when the participants selected the cues and the prediction accuracies achieved when the optimal cue set was provided.

The second major objective of the research was to investigate the information overload phenomenon. Two hypotheses were examined. First, it was hypothesized that prediction accuracy would be maximized after the participants' examination of a small number of cues and then decline as additional cues were examined. Second, it was hypothesized that the participants' confidence in their prediction accuracy would increase as the number of cues used to make the prediction increased.

Both hypotheses were tested using a parametric randomized complete block ANOVA model with multiple comparison procedures to determine if the participants' prediction accuracies (confidence levels) changed when participants based their decisions on two, four, six, and eight cues. The mean prediction accuracies followed the hypothesized pattern. Prediction accuracy increased as the number of cues increased from two to six and then decreased when the number of cues increased to eight. The decrease in prediction accuracy when the number of cues was increased from six to eight was not statistically significant. The first information overload hypothesis, therefore, was not supported statistically.

Like the mean prediction accuracies, the mean confidence levels followed the hypothesized pattern; the participants' confidence in their predictions continued to increase as they received two, four, six, and eight cues. Again, the hypothesis was not supported statistically. The participants' confidence levels significantly increased upon receiving four cues and then remained constant.

In conjunction with the examination of the hypotheses relating to the two major objectives of the study, traditional Brunswick lens model research was replicated. The results of the replication were consistent with those of prior research; the prediction accuracy of the environmental multiple linear regression model was greater than that of both the human decision maker and the multiple linear regression equation representing the decision maker's prediction model. Although many lens model studies have also found the prediction accuracy of the decision maker's model to be greater than that of the human decision maker, in the current study no difference in the prediction accuracy of the participants and their models was found.

Post-test questionnaires were employed to determine if demographic differences, the participants' attitudes toward the study, their degree of understanding of the experimental task and their level of motivation to perform was associated with the participants' performance of the experimental tasks. No relationship between the above factors and the prediction accuracy achieved by the participants was found.

Several other factors were also explored using the post-test questionnaires. One factor examined was the the development and use of heuristics by the participants. From the participants' descriptions of their prediction processes, the one clear strategy employed was anchoring and adjustment. Twelve participants were identified as using this heuristic. There was, however, no clear indication that the participants who used the anchoring and adjustment heuristic had superior prediction accuracies.

Also examined was the participant's intuitive understanding of the information overload phenomenon. Specifically, did the participants recognize that there was a point at which prediction accuracy would not be improved by the use of additional information? Twenty-two of the participants specified a number of cues they believed would maximize prediction accuracy, an indication that these participants did have an intuitive understanding of the information overload phenomenon.

Chapter 6

SUMMARY AND CONCLUSIONS

Both human decision makers and the mathematical models bring advantages to the prediction process. Theoretically, the use of a mathematical model for prediction purposes reduces systematic error first by selecting the optimal set of information cues and second by optimally weighting the cues. Random error, caused by human decision makers' inconsistent use of their own decision models, is also eliminated by the use of a mathematical model.

The human decision maker has two advantages over the mathematical model. First, the mathematical model is a linear model; it does not capture nonlinear relationships, nor does it capture the effects of interactions among the cues. Research has found, however, that human decision makers use information in both a curvilinear and a configural fashion.[135] Second, the human decision maker may bring into the decision making process a wealth of information that cannot be captured by a mathematical model. Empirical research has found that the mathematical model consistently outperforms the human decision maker.[136]

---

[135] Discussed in Chapter Two.

[136] Goldberg, "Man Versus Model of Man", p. 430.

In traditional lens model research, as it pertains to prediction accuracy, a set of optimal cues is selected using a multiple linear regression model. Theses cues are then provided to human decision makers, and their prediction accuracies compared to that of the mathematical model used to select the cues. The results of traditional lens model research have been consistent, the mathematical model outperforms the human decision maker. The traditional lens model research methodology was replicated in this study.

Traditional lens model research has not addressed the question of cue selection. Einhorn has hypothesized that the major factor affecting prediction accuracy is not cue usage but cue selection.[137] To date, only one study has examined the issue of cue selection, and Einhorn's hypothesis was supported.[138] The first major purpose of this research is to determine the effects on prediction accuracy of the human selection of the information cues.

A review of the literature examining information overload indicates that human decision makers may be limited in the number of cues that they can use effectively. Although the results of information overload research are

---

[137]Einhorn, "A synthesis, Accounting and Behavorial Science," p. 167.

[138]Abdel-Khalik and El-Sheshai, "Information Choice and Utilization," pp. 325-342.

mixed, it has been found by some researchers that as the amount of information provided to the human decision maker increases, prediction accuracy first increases and then decreases or becomes constant.[139] The second major purpose of this study is to determine if there is an optimal number of information cues to be provided to the human decision maker. The optimal number of cues is defined as the number of cues with which the human decision maker achieves the highest prediction accuracy.

6.1 SUMMARY OF EMPIRICAL FINDINGS

The following is a discussion of the results of the empirical tests of the hypotheses for information choice, information use, and information overload. A discussion of the possible causes of the results and the conclusions that they suggest are contained in section 6.3 of this chapter.

6.1.1 Information Choice

More than half of the participants achieved a higher prediction accuracy when they selected their own cues than when the "optimal" cue set was provided. Specifically, of the thirty-one participants, sixteen of the participants' environmental models exceeded the prediction accuracy of the environmental model containing the "optimal" cue set,

---

[139] Casey, "Variation in Accounting Information Load", p. 49.

nineteen participants achieved a higher prediction accuracy when they selected their own cues than when provided with the "optimal" cue set, and nineteen of the participants' models achieved a higher prediction accuracy when the participants selected their own cues. The results of the tests of the information choice hypotheses were consistent with the descriptive statistics; there is no difference in the prediction accuracies achieved when the participants selected their own cues and when the "optimal" cue set was provided.

6.1.2 Information Use

In Case Set I, the prediction accuracy of the environmental model was greater than the prediction accuracy of both the participants and their models for twenty-four of the thirty-one participants. Seventeen of the participants achieved a higher prediction accuracy than that of their models. The mean prediction accuracy of the environmental models was $R_{ehI} = 0.7934$, the mean prediction accuracy of the participants was $r_{ahI} = 0.5797$, and the mean prediction accuracy of the participants' models was $r_{mhI} = 0.5590$. To summarize, the prediction accuracy of the environmental models appears to be greater than that of both the participants and their models. The nonparametric multiple comparisons tests support this conclusion at the 0.01 level of significance. There appears to be little difference between the prediction accuracy of the participants and the participants' models. The nonparametric multiple comparisons tests do not conclude that a difference exists.

The descriptive statistics for Case Set II are similar to those of Case Set I. The prediction accuracy of the environmental model was greater than the prediction accuracy of both the participants and their models for thirty of the thirty-one participants. Fourteen of the participants achieved a greater prediction accuracy than that of their models. The prediction accuracy of the environmental model was $R_{emII} = 0.7787$, the mean prediction accuracy of the participants was $r_{amII} = 0.6574$, and the mean prediction accuracy of the participants' models was $r_{mmII} = 0.6454$. The conclusions of the statistical tests are the same as for Case Set I: the prediction accuracy of the environmental model is significantly greater, $a = 0.01$, than the prediction accuracy of both the participants and their models. There is no statistical difference between the median prediction accuracy of the participants and their models.

6.1.3 Information Overload

Multiple comparison procedures were used to test the information overload hypothesis:

H_{a6}: Prediction accuracy (r_{ah}) will be maximized after the decision maker's examination of a small number of cues and then decline as additional cues are examined.

Mean prediction accuracy increased as the number of cues that the participants used to perform the experimental task increased from two, r_{ah2} = 0.6357, to four, r_{ah4} = 0.6894, to six, r_{ah6} = 0.7114, and then decreased when the participants used eight cues, r_{ah8} = 0.6911. There was a significant difference, at the 0.05 level of significance, between the prediction accuracy achieved using two cues and the prediction accuracies using four, six, and eight cues. Although the prediction accuracies indicate a pattern consistent with the information overload hypothesis, the hypothesis was not supported statistically because the decrease in prediction accuracy between r_{ah6} and r_{ah8} was not statistically significant.

A second information overload hypothesis pertaining to the participants' expected levels of confidence was tested:

> Ha_7: The decision maker's confidence in his decision accuracy will increase with the use of additional cues.

The results of the experiment indicate that the participants' mean confidence level in their prediction increased as the number of cues used to make the prediction increased. With two cues, the mean confidence level was 3.9667 on a Likert scale of seven. The mean confidence level increased to 4.4677 when the participants based their decisions on both four cues and six cues. The mean confidence level increased to 4.5312 when the participants based their decisions on eight cues. The increase in mean confidence level was not significant beyond the use of two cues. The conclusion of the multiple comparisons procedures used to test Ha_7 is that there is a significant difference between the participants' mean confidence levels when they based their decisions on two cues and the participants' mean confidence levels when they based their decisions on four, six, and eight cues. There was no significant difference between the participants' mean confidence levels when their predictions were based on four, six, and eight cues.

6.2 CONCLUSIONS

The analysis of the results of the study, presented in Chapter Five, forms the basis for the conclusions to be discussed in this section. These conclusions center on three topics: information choice, information use, and information overload. The conclusions relative to the information overload portion of the study are further divided into discussions of prediction accuracy and confidence level.

6.2.1 Information Choice

The results of the information choice experiment were unexpected. There are, however, several possible causes of the results, and the exploration of these possible causes focuses on three key issues; how "optimal" the optimal cue set, in fact, is; why a high prediction accuracy was achieved by most of the participants; and, if the optimal cue set is optimal, why in Case Set I did seventeen of the cue sets selected by the participants generate a higher prediction accuracy than the optimal cue set.

For several reasons, the discussion of the possible causes of the results will center on the first of the three information choice hypotheses, Ha_3: $R_{emI} - R_{ehI} > 0$. The first reason is that the multiple correlation coefficients of the participants' cue sets, R_{ehI}, and the multiple correlation coefficient of the optimal cue set, R_{emI}, were generated from the same data set. The correlation statistics used to examine the prediction accuracy of the decision makers, r_{amII} and r_{ahI}, and the prediction accuracy of the decision makers' models, r_{mmII} and r_{mhI}, were not generated from the same data sets. Second, the results are not confounded by the abilities of the individual participants. The correlation statistics, R_{emI} and R_{ehI}, are multiple correlation coefficients, the result of multiple linear regression. No human factors affect these statistics. Finally, the results of the tests of the hypotheses for the prediction accuracy of the human decision maker and his model conform to the results of the test of the hypotheses relating to the environmental model.

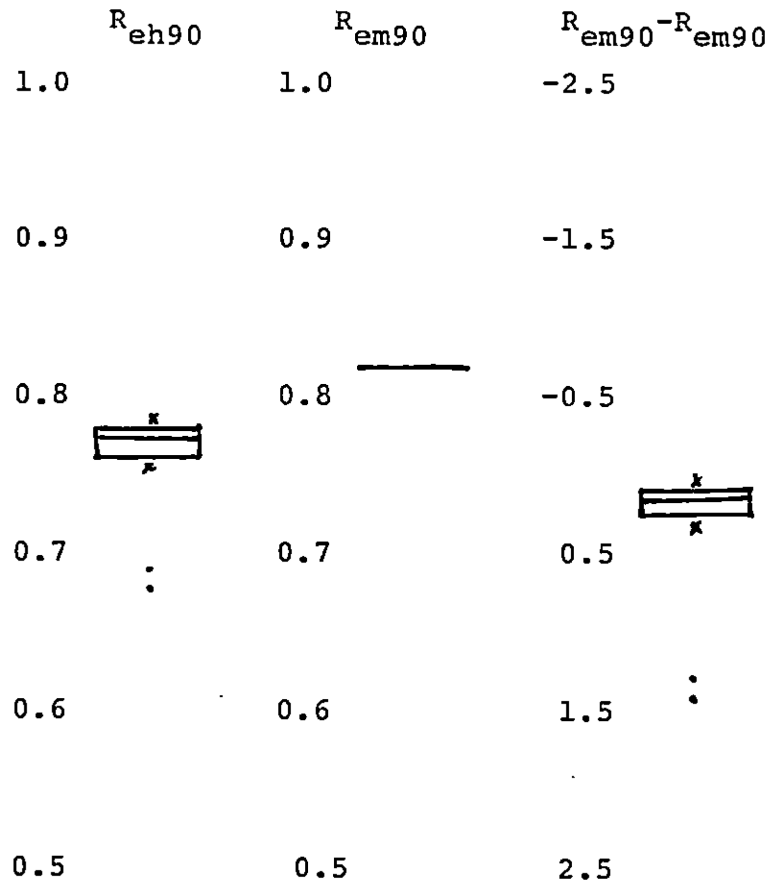
Was the optimal cue set optimal; did the optimal cue set maximize the multiple correlation coefficient? The optimal cue set was selected using the original data set of ninety cases before the data was split into three data sets of thirty cases each. The prediction accuracy achieved by the optimal cue set was $R_{em90} = 0.8092$. The cue selection of each of the participants was regressed on the ninety cases in the original data set. The results are shown in Table 25, p. 196. None of the participants' cue selections achieved a higher prediction accuracy than that of the optimal cue set. The median difference was 0.0215, and the highest prediction accuracy, achieved by participant ten, was $R_{eh} = 0.7998$. A box-plot of the data, Figure 7, p. 197, shows the range of the data, with the exception of the two outliers, to be narrow with a total spread between the first and third quartile of only 0.0126. With the exception of the two outliers, the difference in prediction accuracies achieved is slight. It appears that for the purposes of this research the optimal cue selection was optimal.

Why was a high prediction accuracy achieved by most of the participants? From the examination of Table 25, it appears that there are numerous decision models that will generate a high degree of prediction accuracy. (With the exception of participants twenty-five and three there was little difference in the prediction accuracies achieved by the participants' cue sets and the optimal cue set.)

Table 25

PREDICTION ACCURACIES OF THE ENVIRONMENTAL MODELS:
CASE SET OF NINETY

Participant	R_{eh90}	R_{em90}	$R_{eh90} - R_{em90}$
18	0.7775	0.8092	0.0317
30	0.7786	0.8092	0.0306
18	0.7836	0.8092	0.0256
24	0.7840	0.8092	0.0252
2	0.7881	0.8092	0.0211
4	0.7932	0.8092	0.0160
29	0.7880	0.8092	0.0212
11	0.7877	0.8092	0.0215
15	0.7953	0.8092	0.0139
25	0.6891	0.8092	0.1201
1	0.7922	0.8092	0.0170
21	0.7751	0.8092	0.0341
17	0.7797	0.8092	0.0295
10	0.7998	0.8092	0.0094
13	0.7931	0.8092	0.0161
19	0.7937	0.8092	0.0155
12	0.7887	0.8092	0.0305
20	0.7914	0.8092	0.0178
28	0.7818	0.8092	0.0274
7	0.7894	0.8092	0.0198
23	0.7763	0.8092	0.0329
27	0.7807	0.8092	0.0285
8	0.7836	0.8092	0.0256
31	0.7836	0.8092	0.0256
5	0.7921	0.8092	0.0171
22	0.7955	0.8092	0.0137
26	0.7757	0.8092	0.0335
6	0.7882	0.8092	0.0210
16	0.7784	0.8092	0.0308
12	0.7887	0.8092	0.0353
3	0.6778	0.8092	0.1314
Means:	0.7794	0.8092	0.0298
Medians:	0.7877	0.8092	0.0215
Difference in the medians:	0.0215		



Median	0.7877	0.8092	0.0215
First Quartile	0.7792		0.0301
Third Quartile	0.7918		0.0175
Interquartile Rank	0.0126		0.0126
Step	0.0189		0.0189
Inner Fence	0.7603		0.0490
	0.8107		-0.0014
Outer Fence	0.7414		0.0678
	0.8296		-0.2036

Figure 7

BOX PLOT: PREDICTION ACCURACIES
OF THE ENVIRONMENTAL MODELS
(CASE SET OF NINETY)

Decision makers, therefore, can have individual approaches to the prediction process and select the appropriate cues for their personal decision models with the expectation of achieving a high degree of prediction accuracy. Many participants, for example, appeared to successfully use the anchoring and adjustment heuristic.

Nine participants selected cue number fourteen, "operating income/operating assets," the firm's current year's ROA, as an apparent anchoring point. These nine participants also selected cue eighteen, "operating income growth for the years 19x2 through 19x5," and cue nineteen "operating asset growth for the years 19x2 through 19x5," to determine the trends in both the numerator and the denominator of cue fourteen. All of the nine participants achieved a high degree of prediction accuracy.

Another potential anchor is cue number twenty-nine, "return on assets" for the industry. It was apparently selected as an anchoring point by four participants. These four participants also selected either cue twenty-six, "sales growth during 19x4" for the industry, or cue number twenty-seven, "net income growth during 19x4" for the industry, in an apparent attempt to determine the trend of the numerator of cue twenty-nine. Except for participant number three, all of these participants achieved good prediction accuracy.

A second factor that may contribute to the consistently high degree of prediction accuracy achieved by the participants is the high correlation between many of the most often used cues. The same information is conveyed to different decision makers by different cues. Decision makers, therefore, can develop individual decision models, based on different cue sets, with the expectation of achieving a high degree of prediction accuracy.

Table 26, p. 200, contains the correlations among the cues that were most often selected by the participant decision makers. Cue fourteen, "operating income/operating assets," which was contained in the optimal cue set and was selected by nineteen participants, is highly correlated with cue four, "operating income/sales revenue," which was selected by twelve participants. Cue fourteen is also highly correlated with cue number fifteen, "operating income/total assets," which was selected by nineteen participants. Because there is a high correlation among many of the cues, there are many different cue sets available that will generate a degree of prediction accuracy commensurate with the optimal cue set.

In Case Set I, seventeen of the cue sets selected by the participants generated a higher prediction accuracy than the optimal cue set. This result may be founded in the differences in the data set used to generate the optimal cue list, the full ninety cases, and the data set used by the participants in Case Set I to make their predic-

Table 26

CORRELATION MATRIX OF THE MOST USED CUES

Cue #	4 (12)	9 (19)	14 (19)	15 (12)	18 (20)	19 (17)	20 (7)	25 (7)	26 (8)	29[140] (16)[141]
1										
2										
3	95[142]		71	71						
4			85	86						
5										
6										
7[143]										
8										
9										
10		98								
11										
12		73								
13	73		75	76						
14	85			99						
15	86		99							
16			75	73						
17			76	75						
18							65			
19							91			
20						91				
21										
22										
23										
24										
25										
26										80
27										
28										
29									80	
30									80	94

[140] Cues most often selected by the participants.

[141] Number of participants selecting the cue.

[142] Correlations, the correlation is .95.

[143] Bolded items are the optimal cue set.

tions. The thirty cases for Case Set I were drawn on a random basis from the ninety that were used to generate the optimal cue set. Case Set I is, in effect, a sample thirty firms drawn from a population of ninety. Case Set I can be expected to contain different attributes than the population of ninety because of sampling error.

The results themselves indicate that differences in the data sets exist. The cue set that was optimal for the ninety cases is no longer optimal for the thirty cases in Case Set I. The multiple correlation coefficient using the optimal cue set decreased from 0.8092 for the ninety cases to 0.7925 for Case Set I and to 0.7787 for Case Set II. The distribution of the prediction accuracies of the participants' environmental models is broader for Case Set I, with an interquartile rank of 0.0606, than the distribution of the participants' environmental models for the ninety cases, with an interquartile rank of 0.0126. (Note Figure 7.)

To summarize, the optimal cue set for one data set is not necessarily the optimal cue set for another. The optimal cue set can be determined for a population, but it may not be optimal for a sample of the population because of sampling error. The sample does not necessarily contain all the same attributes of the population.

6.2.2 Information Use

It is concluded from the empirical evidence developed in this study that the prediction accuracy of the environmental models is greater than that of both the participants and their models. It is also concluded that there is no difference between the prediction accuracy of the participants and their models. These conclusions are valid for both case sets.

The prediction accuracy of the human decision maker is a function of:

1. Environmental predictability, R_e, the degree to which the environment can be predicted by the multiple linear regression equation.
2. The matching index, G, the degree to which individual decision makers' models and the environmental model are similar. The environmental model optimally weights the cues. When G is less than unity, decision makers are not giving the optimal weights to the environmental cues, and a systematic bias is created that reduces the prediction accuracy of both the human decision makers and their models.
3. The consistency index, R_s, the consistency with which human decision makers apply their own models to the decision making process. R_s is a random error term, and when it is less than unity decision makers' prediction accuracies suffer as a result of their inconsistent use of their models.
4. Other unknown factors, $r_a - E[r_a]$, which may include

the nonlinear and configural use of the cues by the decision makers and the use of information that was not captured in the regression equation. The existence of other factors is implied by the difference between the prediction accuracies achieved by decision makers and their expected prediction accuracies, $r_a - E[r_a]$. The expected prediction accuracy is computed using the lens model equation:

$$E[r_a] = Re\ G\ R_s$$

The expected prediction accuracy of decision makers, $E[r_a]$, is the prediction accuracy that the decision makers would be expected to achieve considering the degree of environmental predictability, R_e, their accuracy in weighting the cues, G, and the consistency with which they apply their models to the prediction task, R_s.

The above terms are defined in Chapter Two, and the relationships among the correlation statistics R_e, G, r_m, R_e, and r_a are diagrammed in Figure 2, p. 21.

Three of the above factors, the participants' ability to properly weight the cues, measured by the matching index (G); the consistency with which the participants use their own models, measured by the consistency index, (R_s); and other unknown factors, the difference between the participants' prediction accuracy and the expected prediction accuracy of the participants ($r_a - E[r_a]$), provide key insights into the causes of the results achieved by the information use portion of the study. All three factors contribute to the results achieved in Case Set I. First, there was a decrease in prediction accuracy as a result of systematic bias caused by the inability of the participants to weight the cues in an optimal fashion. The median matching index was $G_{hI} = 0.8980$ and the resultant loss in prediction accuracy was $R_{ehI} - r_{mhI} = 0.0883$. (The correlation statistics are reported in Table 27, p. 205.) Second, there was a decrease in prediction accuracy because of random error, caused by the participants inconsistent use of their models; $R_{shI} = 0.9119$. The resultant decrease in prediction accuracy was $r_{mhI} - E[r_{ahI}] = 0.0445$. Finally, other unknown factors caused prediction accuracy to increase by $E[r_{ahI}] - r_{mhI} = 0.0426$.

In Case Set II the consistency index was high; $R_{smII} = 0.9835$, indicating that the participants were applying their own models in a consistent manner. There was little loss of prediction accuracy because of this factor; $r_{amII} - r_{mmII} = 0.0079$. There was also a slight increase in prediction accuracy as a result of unknown factors; $r_{amII} - E[r_{amII}] = 0.0126$. The primary cause of the results achieved in Case Set II was systematic bias created by the participants' inability to weight the cues in an optimal manner. The median matching index, G_{mII}, was 0.8812. If the participants had weighted the cues in the optimal

Table 27

INFORMATION USE:
SUMMARY OF STATISTICS

Correlation Statistics: Medians

	R_e	G	r_m	R_s	r_a	$E[r_a]$
Case Set I	0.7954	0.8980	0.7071	0.9119	0.7052	0.6626
Case Set II	0.7787	0.8812	0.6862	0.9835	0.6783	0.6657

Hypotheses Tested, Type of Test, and P-values

Hypotheses:	Test	P-value
Ha: R_{smII} R_{shI}	Sign	0.0003
Ha: r_{ahI} $E[r_{ahI}]$	Sigh	0.0239

manner, the matching index would have been one; and the prediction accuracy of the participants' models would exactly equal that of the environmental model. The resulting loss in prediction accuracy was $R_{emII} - r_{mmII} = 0.0925$.

Although the conclusions of the statistical tests of the prediction accuracies of both cases were consistent, the prediction accuracy of the environmental models is greater than that of both the participants and their models and there is no difference between the prediction accuracy of the participants and their models. There are two key differences in the causes of the results. First, the median consistency index was less for Case Set I, $R_{shI} = 0.9119$, than for Case Set II, $R_{smII} = 0.9835$. Twenty-five of the thirtyone participants had a higher consistency index in Case Set II. The difference is statistically significant, with a p-value of $p = 0.0003$. The more consistent use of their own models by the participants may be an indication that when the participants select their own cues, they lack confidence in their models.

A second difference is that the other unknown factors, $r_a - E[r_a]$, caused a meaningful increase in prediction accuracy in Case Set I. Twenty-seven of the thirty-one participants had an increase in prediction accuracy as a result of other factors. The difference is statistically significant, $p < 0.0002$. This increase may indicate that when the participants select their own models, they are able to more effectively interrelate nonquantifiable cues and/or use the cues selected in a nonlinear and configural fashion thereby generating a meaningful increase in prediction accuracy.

6.2.3 Information Overload

Two elements relative to the information overload phenomenon were examined. The participants' achieved prediction accuracies and their reported confidence levels in their decisions were analyzed.

6.2.3.1 Prediction Accuracy

Although the prediction accuracies followed the predicted pattern (Table 14, p. 153), the expectation that prediction accuracy would first increase and then decrease as the decision maker received additional cues was not supported statistically. Statistically, there was a significant increase in prediction accuracy when the number of cues used by the participant increased from two to six. The decrease in prediction accuracy between six and eight cues was not statistically significant. Despite the results of the statistical tests, it is concluded that information overload did, in fact, occur. There are three reasons for this conclusion. First, there was a decrease in median prediction accuracy when the number of cues increased from six to eight. Second, when the data was disaggregated, twenty-five of the thirty-one participants experienced information overload, a decrease in prediction accuracy taking place by the time the participant had used eight cues. Finally, causality is clearly evident.

The four components of prediction accuracy discussed in section 6.2.2 must be examined to determine the causes of the results achieved. Three of the components, environmental predictability (R_e), the participants' ability to properly weight the cues (G), and the consistency with which the participants use their own models (R_s), explain the changes in the participants' prediction accuracies. The fourth component, other factors ($r_a - E[r_a]$), does not appear to be a causal component because there is no trend and the changes in prediction accuracy as a result of this component are small. This conclusion is confirmed by the graph of the means of the participants' models (Mean Expected Prediction Accuracy, Table 28, p. 209). The curve is similar to that of the participants' mean prediction accuracies (Table 28). Prediction accuracy increases as the participant uses four and six cues and decreases when eight cues are used. (Table 29, p. 210 and Figure 8, p. 211, provide representations of the four components of the participants' prediction accuracies.)

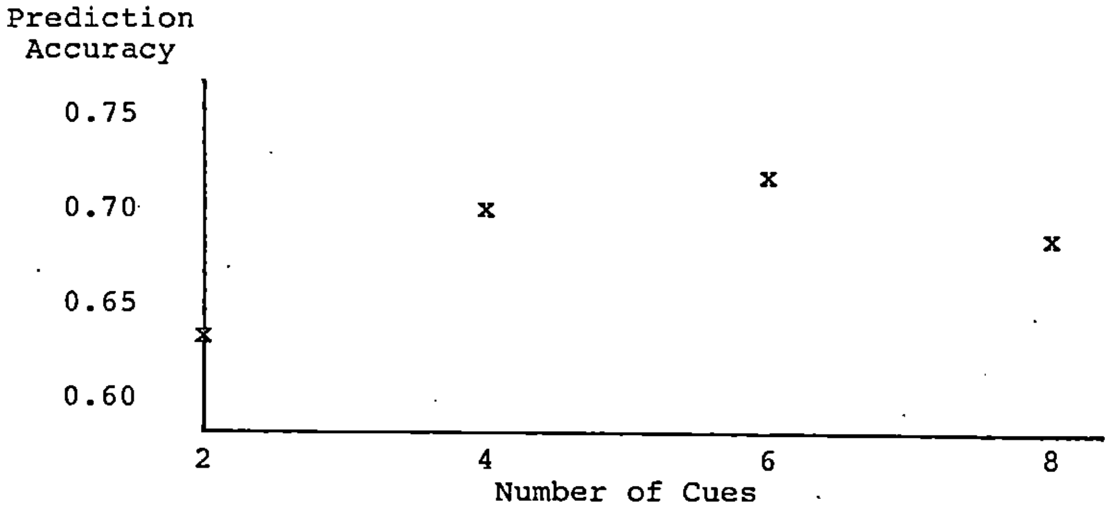
The environmental model is a multiple linear regression equation, and as a result, the multiple correlation coefficient, Re, will increase as cues are added to the model. The mean change in the multiple correlation coefficient was $R_{e8} - R_{e2} = 0.1309$, with the major portion

Table 28

INFORMATION OVERLOAD:
PREDICTION ACCURACY AND EXPECTED PREDICTION ACCURACY

Number of Cues	Mean Prediction Accuracy	Mean Expected Prediction Accuracy
2	0.6357	0.6580
4	0.6911	0.7029
6	0.7114	0.7429
8	0.6894	0.7118

Mean Prediction Accuracy



Mean Expected Prediction Accuracy

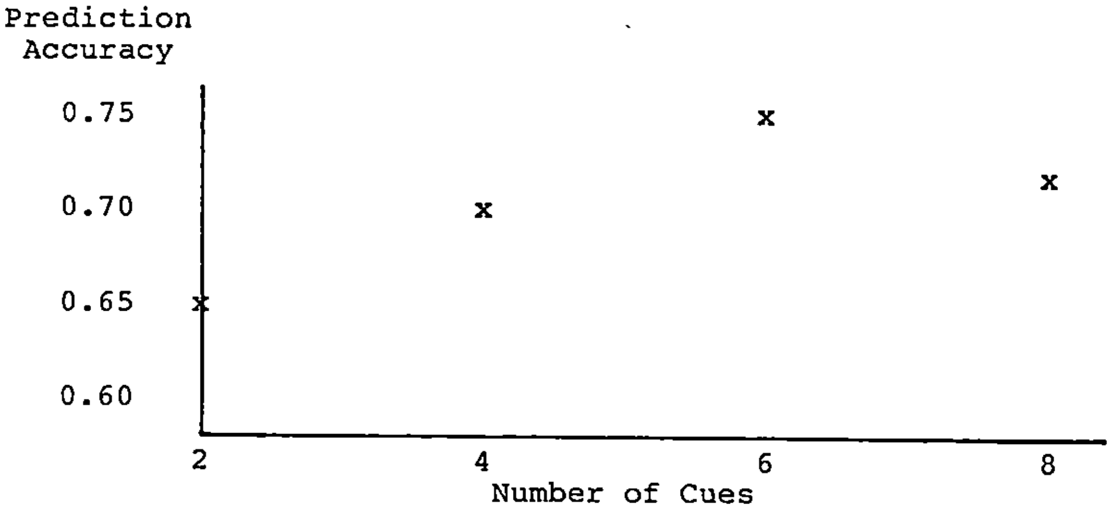


Table 29

INFORMATION OVERLOAD:
COMPONENTS OF PREDICTION ACCURACY

Correlation	Number of Cues 2	4	6	8
R_e	0.7460	0.8224	0.8456	0.8769
G	0.8645	0.9089	0.9253	0.8531
R_s	0.8868	0.9067	0.9497	0.9461
$ra - E[r_a]$	0.0223	0.0118	0.0315	0.0224

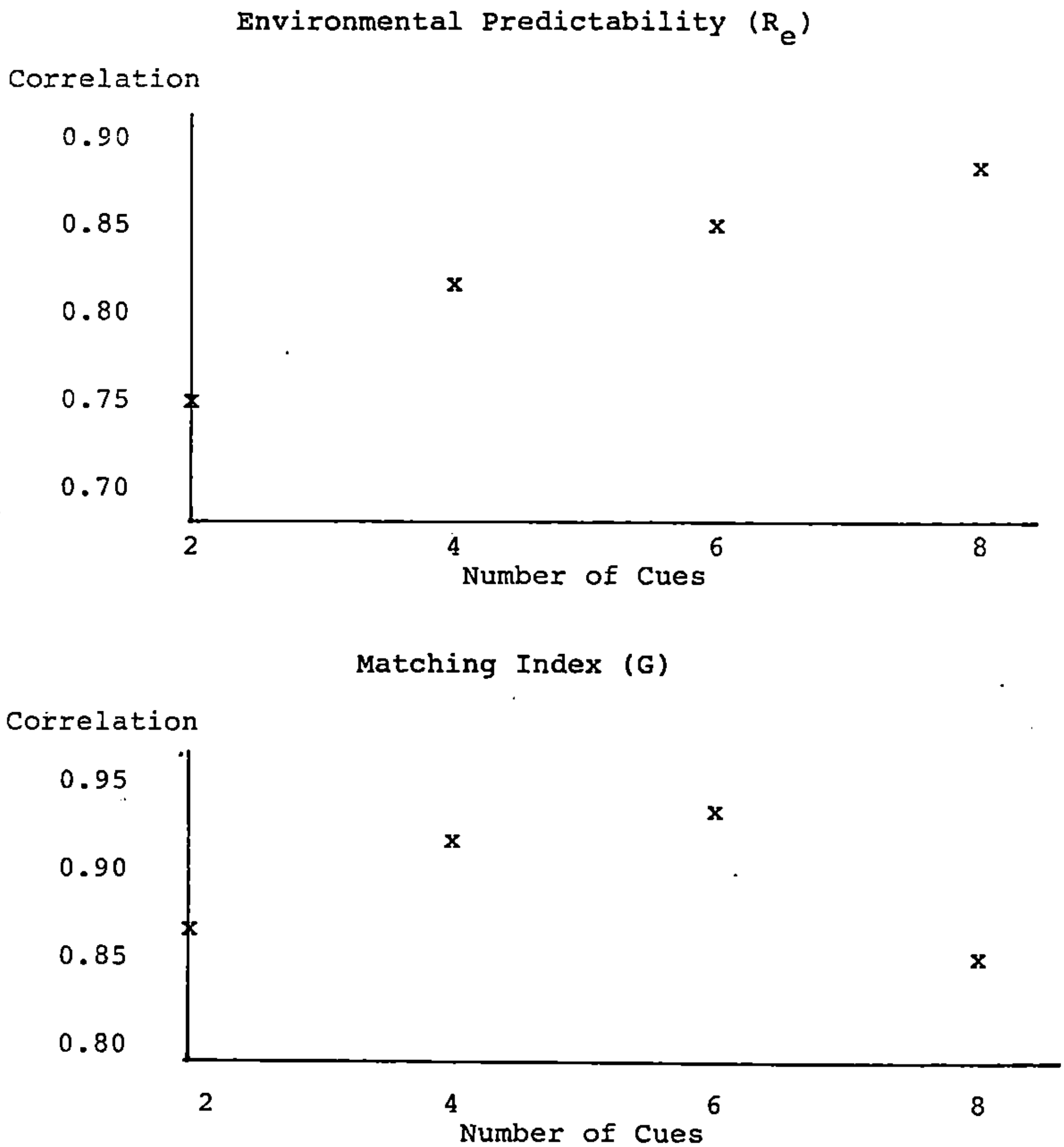


Figure 8

INFORMATION OVERLOAD:
COMPONENTS OF PREDICTION ACCURACY

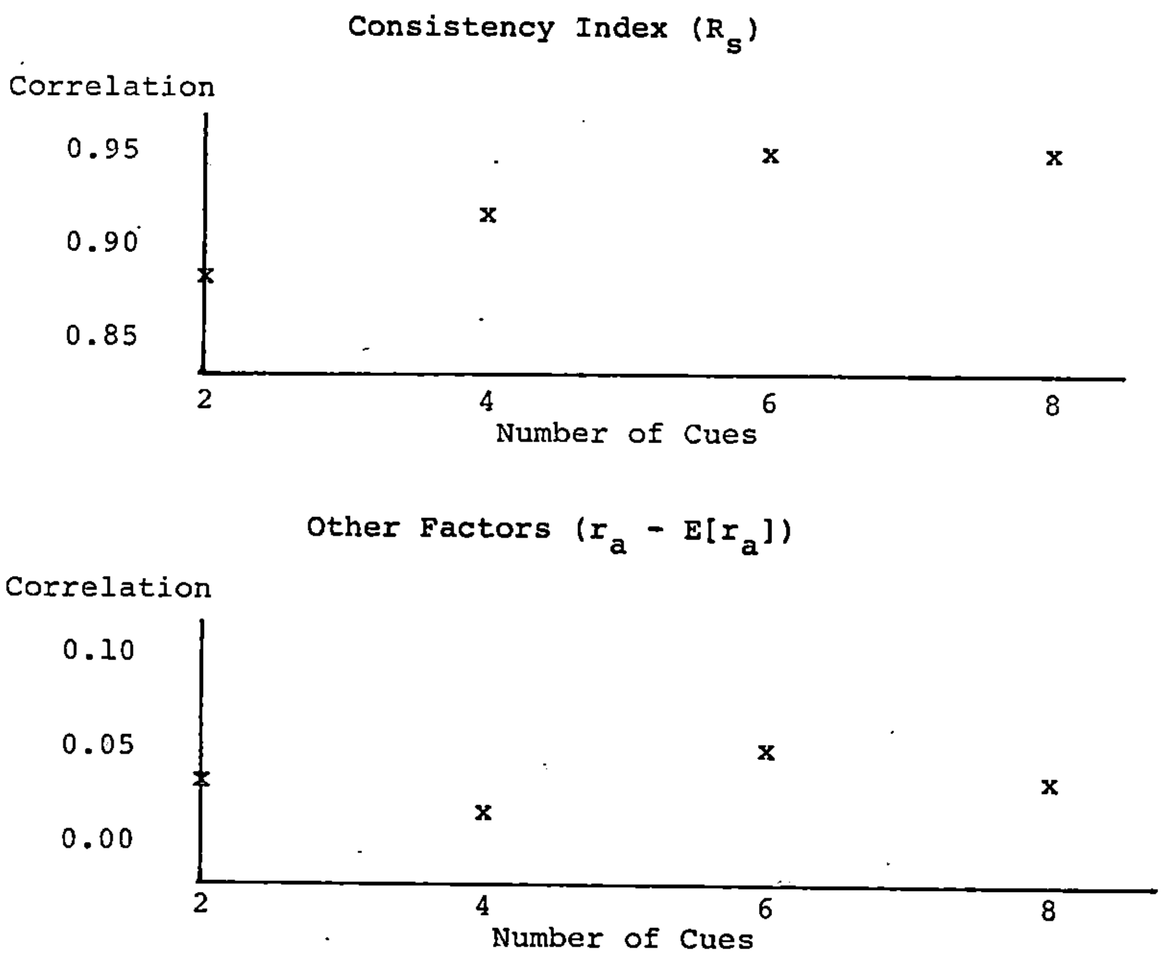


Figure 8 (Continued)

of the change taking place when the number of cues changed from two to four, $R_{e4} - R_{e2} = 0.0764$. This component alone causes a significant increase in prediction accuracy. The minimum significant difference for the Tukey's multiple comparison procedure was 0.0736 at the 0.05 level of significance, and the increase in prediction accuracy was 0.0832 (r_{a2}, 0.6357, multiplied by change in the multiple correlation coefficient, 0.1309).

The cause of the decrease in the participants' prediction accuracy, information overload, is the decrease in the participants' ability to properly weight the cues as the number of cues increases. The mean matching index increased from 0.8645 for two cues to 0.9253 for six cues and then decreased sharply to .8531 for eight cues. The matching index decreased for thirty of thirty-one participants as the number of cues increased. The mean decrease in the matching index that took place when eight cues were used was greater than the increase in prediction accuracy caused by the increase in the multiple correlation coefficient. The decrease in prediction accuracy caused by the participants' inability to properly weight the cues was not significant.

As a result of the increase in the consistency index, $R_{s8} - R_{s2} = 0.0296$, there was a slight increase in prediction accuracy of 0.0189. The index was stable for cues two and four, increased by 0.0430 at six cues and then decreased slightly ($R_{s8} - R_{s6} = 0.0036$). The consistency index does not appear to be a major component of the participants' prediction accuracies.

The change in the participants' prediction accuracies as they received new information is a result of the interaction of three of the components of prediction accuracy. First, prediction accuracy significantly increased as a result of increased environmental predictability. Second, prediction accuracy first increased and then decreased as a result of the participants' inability to properly weight the cues as they received additional cues. The decrease in the matching index appears to be the major cause of information overload. Finally, prediction accuracy decreased slightly when the participants received eight cues as a result of the participants' inconsistent use of their own models. (These factors are graphed in Figure 8.)

6.2.3.2 Confidence

The participants' confidence levels followed the predicted pattern; however, the expectation that the decision makers' confidence in their decisions would increase as they received additional cues was not statistically supported. Statistically, the participants' mean confidence levels increased upon the participants' receiving four cues and then remained constant.

6.3 LIMITATIONS

The use of students as surrogates for actual decision makers limits the research in two ways. First, an advantage of the human decision maker is that he brings into the prediction process background information that while not mathematically quantifiable, may contribute to prediction accuracy. The student participant does not bring the depth and breadth of background into the decision making situation that an actual decision maker would.

A second limitation of using students as surrogates is that the student may not have the forecasting skills of actual decision makers. In order to mitigate this limitation, the participants were carefully selected from the Masters of Accountancy and the MBA programs at Virginia Tech to insure, as much as possible, that they were familiar with ROA and the factors causing its change.

After examining numerous studies that compared the responses of students and actual decision makers, Ashton and Kramer found that "studies which have focused on *decision making* have found considerable similarities in the decisions and the apparent underlying information processing behavior of student and nonstudent groups."[144] While it is recognized that the use of students as surrogates for

---

[144] Robert H. Ashton and Sandra S. Kramer, "Students as Surrogates in Behavioral Accounting Research: Some Evidence," *Journal of Accounting Research*, (Spring 1980), p. 1.

actual decision makers limits the ability to generalize the research results, students are acceptable as participants because the focus of the research is on information processing and its relationship to prediction accuracy.

Another possible limitation is that there is not a perfect pairing of the data for the tests of all three hypotheses. A perfect pairing is achieved for the first hypothesis; the prediction accuracy of the participants' cue sets was compared with the prediction accuracy of the optimal set of cues for each participant using the same data set, Case Set I. For the second hypothesis, which tests the prediction accuracy of the decision maker, and for the third hypothesis, which tests the prediction accuracy of the decision makers' models, a complete pairing was not achieved. The differences between participants were controlled for by having each participant perform the experiment twice, first using his own cue selection, then using the optimal cue set. However, each experiment was performed using different data sets, Case Set I and Case Set II. Two data sets were used because if the participants performed the experiment twice on the same data set, there would be a potential for learning.

It is believed the use of a different set of cases for each experiment did not affect the results of the tests of the hypotheses. First, Case Set I and Case Set II were randomly selected from the ninety original cases. Second, there is no material difference in the p-values of the sign test when the first hypothesis is tested using Case Set I only, and when Case Set I and Case Set II are used. Using Case Set I, eighteen of the participants' environmental models generated a higher prediction accuracy than the environmental model using the optimal cue set, with a p-value for the sign test of $p = 0.85720$. Using Case Set II, sixteen of the participants' environmental models generated a higher prediction accuracy than the environmental model using the optimal cue set, with a p-value for the sign test of $p = 0.81541$.

6.4 SUGGESTIONS FOR FUTURE RESEARCH

The information choice portion of this study found that human decision makers are capable of making highly accurate predictions when selecting their own cues. This finding appears to be the result of several factors:

1. There is a high correlation among the information cues upon which financial decision makers base their predictions. There are, therefore, potentially a large number of models that could be highly predictive.
2. The participants appeared to have a greater confidence in the models that they selected themselves. Twenty-five of the thirty-one participants had a higher consistency index when they selected their own cues.

3. There was a meaningful increase in the participants' prediction accuracies as a result of unknown factors (r_a - $E[r_a]$). This increase may indicate that when the participants select their own models, they are better able to more effectively interrelate nonquantifiable cues and/or use the cues selected in a nonlinear or configural fashion.
4. An environmental model developed using one data set may not be the best model for other data sets. The predictability of regression equations is particularly affected by the presence of influential subsets of data and outliers.

This research has shown that human decision makers are capable of a high degree of prediction accuracy. Research is needed to identify and define the advantages that human decision makers bring to the decision process and to define the proper role of both the human and the mathematical model in the decision process so that the advantages of both can best be used. Studies are needed that determine the human decision maker's ability to adapt to anomolies in data sets. The effects on prediction accuracy of the number of cues, cue validities, correlations among cues, and the ability of the decison maker to use nonquantifiable cues need to be examined.

The conclusion of the information overload portion of this study was that prediction accuracy was maximized when the participants based their decisions on six cues and then deteriorated with the addition of new cues. In general, the results of information overload studies have been mixed. More research is needed using actual decision makers in actual decision making situations, allowing the participants to select their own cues and the number of cues that they would like to receive in the cases. The relationships between information overload and cue validities and the correlation among cues might also be examined more closely.

BIBLIOGRAPHY

Abdel-khalik, R. A. "The Effect of Aggregation of Accounting Reports on the Quality of the Lending Decision: An Empirical Investigation." Supplement to Journal of Accounting Research (1973): 104-138.

Abdel-khalik, R. A. and El-Sheshai, K. M. "Information Choice and Utilization in an Experiment on Default Prediction." Journal of Accounting Research (Autumn 1980): 325-342.

Altman, E. and McGough, T. "Evaluation of a Company as a Going Concern." Journal of Accountancy (December 1974): 50-57.

American Accounting Association, A Statement of Basic Accounting Theory. Sarasota, Florida: AAA, 1966.

Ashton, R. H. Human Information Processing in Accounting. Sarasota, Florida: American Accounting Association, 1981.

Ashton, R. H. "The Predictive Ability Criterion and User Prediction Models." The Accounting Review (October 1974): 719-732.

Ashton, R.H. and Kramer, S. "Students as Surrogates in Behavioral Accounting Research: Some Evidence." Journal of Accounting Research (Spring 1980): 1-16.

Barefield, R. M. "The Effect of Aggregation on Decision Making Success: A Laboratory Study." Journal of Accounting Research (Autumn 1972): 229-242.

Belsley, David A.; Kun, Edwin; and Welsh, Ray E. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New York: John Wiley & Sons, 1980.

Beaver, W. H. "Financial Ratios as Predictors of Business Failure." Supplement to Journal of Accounting Research (1976): 196-206.

Beaver, W. H. "Financial Ratios and Predictors of Failure." Supplement to Journal of Accounting Research (1966): 71-111.

Brunswick, E. The Conceptual Framework of Psychology. Chicago: University of Chicago Press, 1952.

Burger, A. D. and Webster, S. K. "The Management Accountant Looks at EPS vs. ROA: Conflict in Measuring Performance." Management Accounting (August 1978): 19-23.

Casey, C. J. "The Usefulness of Accounting Ratios for Subjects' Predictions of Corporate Failure: Replication and Extensions." Journal of Accounting Research (Autumn 1980): 603-613.

Casey, C.J. "Variation in Accounting Information Load: The Effect on Loan Officers' Predictions of Bankruptcy." The Accounting Review (January 1980): 36-49.

Committee on Accounting Theory Construction and Verification. "Report of the Committee on Accounting Theory Construction and Verification." Supplement to The Accounting Review (1971): 59-73.

Dawes, R. M. "A Case Study of Graduate Admissions: Application of Three Principles of Human Decision Making." American Psychologist (February 1971): 180-188.

Dawes, R. M. and Corrigan, B. "Linear Models in Decision Making." Psychological Bulletin (February 1974): 95-106.

Devine, C. T. "Research Methodology and Accounting Theory Formation." The Accounting Review (July 1960): 387-399.

Dudycha, L. W. and Naylor, J. C. "Characteristics of the Human Inference Process in Complex Choice Behavior Situations." Organizational Behavior and Human Performance (September 1966): 110-128.

Ebert, R. J. and Kruse, T. E. "Bootstrapping the Security Analyst." Journal of Applied Psychology (February 1978): 110-119.

Einhorn, H. J. "Expert Measurement and Mechanical Combination." Organizational Behavior and Human Performance (February 1972): 86-106.

Einhorn, H. J. "A Synthesis, Accounting and Behavioral Science." Supplement to Journal of Accounting Research (1976): 167-179.

Einhorn, H. J. "Use of Nonlinear, Noncompensatory Models as a Function of Task and Amount of Information." Organizational Behavior and Human Performance (January 1971):

1-27.

Financial Accounting Standards Board. Statement on Financial Accounting Concepts No. 1. Stamford, Connecticut: FASB, 1978.

Goldberg, L. R. "Diagnosticians vs. Diagnostic Signs: The Diagnosis of Psychosis vs. Neurosis from MMPI." Psychological Monographs: General and Applied (1965): 1-28.

Goldberg, L. R. "Man Versus Model of the Man: A Rationale, Plus Some Evidence, for a Method of Improving on Clinical Inferences." Psychological Bulletin (June 1970): 422-432.

Goldberg, L. R. "Simple Models or Simple Process? Some Research on Clinical Judgments." American Psychologist (July 1968): 483-496.

Hammond, K. R. and Summers, D. A. "Cognitive Control." Psychological Review (January 1972): 58-67.

Iman, R. L. and Conover, W. J. Modern Business Statistics. New York: John Wiley & Sons, 1983.

Kida, T. "An Investigation into Auditors' Continuity and Related Qualification Judgments." Journal of Accounting Research (Autumn 1980): 506-523.

Kleinbaum, D. G. and Kupper, L. L. Applied Regression Analysis and Other Multivariable Methods. North Scituate: Duxbury Press, 1978.

Libby, R. Accounting and Human Information Processing: Theory and Applications. Englewood Cliffs, New Jersey: Prentice-Hall, 1981.

Libby, R. "Accounting Ratios and the Prediction of Failure: Some Behavioral Evidence." Journal of Accounting Research (Spring 1975): 150-161.

Libby, R. "The Use of Simulated Decision Makers in Information Evaluation." Accounting Review (July 1975): 475-489.

Libby, R. "Man Versus Model of Man: Need for a Nonlinear Model." Organizational Behavior and Human Performance (June 1976): 23-26.

Libby, R. "Man Versus Model of Man: Some Conflicting Evidence." Organization Behavior and Human Performance (July 1976): 23-26.

Miller, G. A. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information." The Psychological Review (March 1956): 81-97.

Montgomery, Douglas C. and Peck, Elizabeth A. Introduction to Regression Analysis. New York: John Wiley & Sons, 1982.

Neter, John; Wasserman, William; and Kutner, Micheal H. Applied Linear Statistical Models. Homewood, Illinois: Richard D. Irwin, Inc., 1985.

Oscamp, S. "Overconfidence in Case-study Judgments." Journal of Consulting Psychology (June 1965): 261-265.

Pollack, I. "The Assimilation of Sequentially Encoded Auditory Displays." American Journal of Psychology (1953): 745-749.

Pollack, I. and Flicks, L. "Information of Elementary Multi-dimensional Auditory Displays." Journal of Accoustics Society of America (1954): 155-158.

Siegel, J. G. "How Management Can Make Significant Use of Financial Analysis to Evaluate the Health of A Firm." The National Public Accountant (February 1982): 25-29.

Slovic, P. and Litchenstein, S. "Comparison of Bayesian and Regression Approaches to the Study of Information Processing Judgment." Organizational Behavior and Human Performance (1971): 649-743.

Wiggins, N. and Kohen, E. S. "Man vs. Model of Man Revisited: The Forecasting of Graduate School Success." Journal of Personality and Social Psychology (July 1971): 100-106.

Wright, W. F. "Financial Information Processing Models: An Empirical Study." The Accounting Review (July 1977): 676-689.

Wright, W. F. "Properties of Judgment Models in a Financial Setting." Organizational Behavior and Human Performance (February 1979): 73-85.

Wright, W. F. "Self-insight into the Cognitive Processing of Financial Information." Accounting Organizations and Society (1977): 323-331.

Yntema, D. B. and Torgerson, W. S. "Man-Computer Cooperation in Decisions Requiring Common Sense." IRE Transactions of the Professional Group on Human Factors in Electronics (1961): 20-26.

Zimmer, I. "A Comparison of the Prediction of Loan Officers and Their Linear-Additive Models." Organizational Behavior and Human Performance (February 1981): 69-73.

Zimmer, I. "A Lens Study of the Prediction of Corporate Failure by Bank Loan Officers." Journal of Accounting Research (Autumn 1980): 629-636.

APPENDIX A

Information Choice/Use

DIRECTIONS FOR THE CASES

Reward structure: Two groups of thirty cases each will be displayed on the terminal. Each case represents an actual firm. You will be requested to predict an individual firm's future ROA based on the information provided and to record your confidence in the prediction.

For taking part in the study each participant will receive ten dollars. The ten dollar remuneration will be paid upon the completion of the cases and a post test questionnaire. In addition, a twenty-five dollar award, based on prediction accuracy, will be paid to the participant who achieves the highest overall prediction accuracy. The twenty-five dollar awards will be mailed on June 5, 1984.

Purpose: The purpose of accounting data is to provide information that is useful in predicting future events upon which financial decisions are based. This study is designed to examine several factors that may affect the prediction accuracy and thereby decision accuracy. One of these factors is the decision maker's ability to select information cues that are related to the event that is being predicted. Other factors include the decision maker's ability to appropriately weight the cues selected and to use the cues selected in a consistent fashion to achieve an accurate prediction.

<u>The task</u>: Assume that it is January 1, 19x5 and you are attempting to predict the return on assets (ROA) for thirty manufacturing firms at December 31, 19x5. To predict the December 31. 19x5 ROA ratios for the thirty firms, you have been provided a "shopping list" of eight financial ratios, the CUE LIST, that have been found to be predictive of a firm's future ROA.

The general procedure for completing the first set of cases, Group I, is to select from the "shopping list" the six ratios you believe to be the best predictors of the December 31, 19x5 ROA ratios for the thirty firms. The ratios selected will be automatically generated for each of the thirty firms by the computer and displayed at the terminal in a case format. The cases have been created using actual manufacturing firms and each of the thirty firms is an individual case.

To complete each case simply record your prediction of the firm's December 31, 19x5 ROA and the level of confidence that you have in your prediction. After the first case has been completed, the second case will be displayed on the screen.

ROA is defined as operating income divided by operating assets. Your confidence level is based on a seven point scale. The scale is

1	2	3	4	5	6	7
No Confidence						Extreme Confidence

A four would represent an average level of confidence in your prediction. The numbers to the left of four indicate a less than average level of confidence in your prediction and the numbers to the right of four indicate an above average level of confidence in your prediction.

After completing the Group I cases, continue on to complete the Group II cases. The procedure is the same as in the Group I cases except that the optimal set of six cues has been provided for you. Again, record your estimate of the firm's ROA and your level of confidence.

Specific Procedures: The specific procedures for completing the study will be given verbally by the experimenter.

Hints:

1. Only two control keys are needed to complete the study, the tab key and the "enter" key.
2. Study the CUE LIST carefully before beginning the study.
3. Keep track of the cues selected.
4. Please do not hesitate to ask questions at any time.

5. After completing the study notify the experimenter. You will be given a short questionnaire to complete.

Information Choice/Use

CUE LIST

Firm Specific Cues

Operating margins:

1. Operating expenses/sales revenue.
2. Working capital provided by operations/sales revenue.
3. Operating income before depreciation/sales revenue.
4. Operating income/sales revenue.
5. Net income/sales revenue.
6. Effective tax rate: Income taxes/net income before taxes.

Turnover ratios:

7. COGS/average inventory.
8. Working capital/sales revenue.
9. Sales/operating assets.
10. Sales/total assets.
11. Sales/common stockholder's equity.
12. Sales/total stockholder's equity.

Rates of return:

13. Working capital provided by operations/total assets.
14. Operating income/operating assets.
15. Operating income/total assets.
16. Net income minus preferred dividends/common stockholder's equity.
17. Net income/total stockholder's equity.

Trend ratios:

18. Operating income growth for the years 19x2 through 19x4.
19. Operating asset growth for the years 19x2 through 19x4.
20. Sales growth for the years 19x2 through 19x4.
21. Capital spending growth for the years 19x2 through 19x4.
22. Growth in sales/employees ratio for the years 19x2 through 19x4.

Other firm ratios:

23. Current ratio.
24. Total debt/total assets.
25. Market share.

Industry Cues

26. Sales growth during 19x4.
27. Net income growth during 19x4.
28. Operating profit margin.
29. Return on assets.
30. Return on equity.

Information Choice/Use

QUESTIONNAIRE

NAME: ____________________

SS#: ____________________

Please take a few additional minutes to answer the following questions to the best of your ability.

1. Age? Years: ______

2. Have you ever had a course in financial statement analysis? (Circle one answer.)
   Yes
   No

3. Do you have job experience in professional accounting or finance? (Circle one answer.)
   Yes
   No

4. In what field do you expect your entry level position to be upon graduation? (Circle one answer.)
   a. Accounting
   b. Finance
   c. General business
   d. Management science
   e. Marketing
   f. Other (explain)

5. For each item circle the number that best indicates your general reaction to the study.

Boring						Interesting
1	2	3	4	5	6	7
Fatiguing						**Stimulating**
1	2	3	4	5	6	7
Unrealistic						**Realistic**
1	2	3	4	5	6	7

6. Did you find the study to be: (Circle one answer.)
   a. Educational
   b. Not educational

7. Did the potential for reward simulate you to perform to the best of your abilities. (Circle one answer.)
   Yes
   No

8. Did you develop a strategy before you selected your cue set for the Group I cases? (Circle one answer.)
   Yes
   No

   If yes, please describe the strategy.

9. If the answer to question 8 above is no, did you develop a strategy at any time during the experiment? (Circle one answer.)
   Yes
   No

   If yes, at what point did you develop the strategy?

   Please describe the strategy.

10. Would you volunteer to participate in such an exercise again? (Circle one answer.)
    Yes
    No

11. If you could give one piece of advice to the experimenter, what would it be?

12. Would you like a summary of your prediction accuracy? (Circle one answer.)
    Yes
    No

PLEASE DO NOT DISCUSS THIS STUDY WITH OTHERS. Don't forget to pick up your ten dollars on the way out. Thanks.

Information Overload

DIRECTIONS FOR THE CASES

Reward structure: A group of cases, each representing an actual firm, will be displayed on the terminal. You will be requested to predict each individual firm's future ROA based on the information provided and to record your confidence in the prediction.

For taking part in the study each participant will receive ten dollars. The ten dollar remuneration will be paid upon the completion of the cases and a post test questionnaire. In addition, two, twenty-five dollar awards, based on prediction accuracy, will be paid. The first award will be paid to the participant who achieves the highest overall prediction accuracy. The second will be paid to the subject who achieves the highest correlation between accuracy and confidence level. The twenty-five dollar awards will be mailed on June 5, 1984.

Purpose: The purpose of accounting data is to provide information that is useful in predicting future events upon which financial decisions are based. This study is designed to examine several factors that may affect the prediction accuracy and thereby decision accuracy. One of these factors is the number of information cues used to make the prediction. Other factors include the decision maker's ability to select information cues that are related to the event that is being predicted and his ability to appropriately weight the cues selected to achieve an accurate prediction.

<u>The</u> <u>task</u>: Assume that it is January 1, 19x5 and you are attempting to predict the return on assets (ROA) for thirty manufacturing firms at December 31, 19x5. To predict the December 31. 19x5 ROA ratios for the thirty firms, you have been provided a "shopping list" of eight financial ratios, the CUE LIST, that have been found to be predictive of a firm's future ROA.

The general procedure for this study is to select from the "shopping list" the two ratios you believe to be the best predictors of the December 31, 19x5 ROA ratios for the thirty firms. The ratios selected will be automatically generated for each of the thirty firms by the computer and displayed at the terminal in a case format. The cases have been created using actual manufacturing firms and each of the thirty firms is an individual case.

To complete each case simply record your prediction of the firm's December 31, 19x5 ROA and the level of confidence that you have in your prediction. After the first case has been completed, the second case will be displayed on the screen. It should take only a few minutes to complete the thirty cases.

ROA is defined as operating income divided by operating assets. Your confidence level is based on a seven point scale. The scale is

1	2	3	4	5	6	7
No Confidence						Extreme Confidence

A four would represent an average level of confidence in your prediction. The numbers to the left of four indicate a less than average level of confidence in your prediction and the numbers to the right of four indicate an above average level of confidence in your prediction.

After completing all thirty cases, select two additional cues and repeat the cases recording your revised estimates of the firm's ROA and your level of confidence based on the new information. The procedure will be repeated until all eight cues have been used in the prediction process and the initial prediction and confidence level have been revised three times.

Specific Procedures: The specific procedures for completing the study will be given verbally by the experimenter.

Hints:

1. Only two control keys are needed to complete the study, the tab key and the "enter" key.
2. Study the CUE LIST carefully before beginning the study.
3. Keep track of the cues selected.

4. If you have any questions, please call on the experimenter.
5. After completing the study notify the experimenter. You will be given a short questionnaire to complete.

Information Overload

CUE LIST

1. Net income/stockholders equity for 19x4.
2. Inventory/COGS for 19x4.
3. Operating income growth for the years 19x2 through 19x4.
4. Sales growth for the years 19x2 through 19x4.
5. Capital spending growth for the years 19x2 through 19x4.
6. ROA for 19x4: Operating income/operating assets.
7.* Industry ROA for 19x4: Operating income/ operating assets.
8.* Industry sales growth for 19x4.

* All ratios are for the firm except numbers seven and eight, which are industry ratios.

Information Overload

QUESTIONNAIRE

NAME: ____________________

SS#: ____________________

Please take a few additional minutes to answer the following questions to the best of your ability.

1. Age? Years: ______

2. Have you ever had a course in financial statement analysis? (Circle one answer.)
   Yes
   No

3. Do you have job experience in professional accounting or finance? (Circle one answer.)
   Yes
   No

4. In what field do you expect your entry level position to be upon graduation? (Circle one answer.)
   a. Accounting
   b. Finance
   c. General business
   d. Management science
   e. Marketing
   f. Other (explain)

5. For each item circle the number that best indicates your general reaction to the study.

Boring						Interesting
1	2	3	4	5	6	7

Fatiguing						Stimulating
1	2	3	4	5	6	7

Unrealistic						Realistic
1	2	3	4	5	6	7

6. Did you find the study to be: (Circle one answer.)
   a. Educational
   b. Not educational

7. Did the potential for reward simulate you to perform to the best of your abilities. (Circle one answer.)
   Yes
   No

8. At what point during the experiment did you begin to experience fatigue? (Circle one answer.)
   a. After 2 cues
   b. After 4 cues
   c. After 6 cues
   d. After 8 cues
   e. I did not find the study fatiguing.

9. Did you develop a strategy before you selected the first two cues? (Circle one answer.)
   Yes
   No

   If yes, please describe the strategy.

10. If the answer to question 9 above is no, did you develop a strategy at any time during the experiment? (Circle one answer.)
    Yes
    No

    If yes, at what point did you develop the strategy?

    Please describe the strategy.

11. Please indicate the number of cues that you believe are required to adequately predict ROA. (Circle one answer.)
    a. Two cues
    b. Four cues
    c. Six cues
    d. Eight cues
    e. More than eight cues

    Do you believe that your prediction accuracy improved, or would you improve, as a result of continuing the experiment beyond the number of cues indicated above? (Circle one answer.)
    Yes
    No

12. Would you volunteer to participate in such an exercise again? (Circle one answer.)
    Yes
    No

13. If you could give one piece of advice to the experimenter, what would it be?

14. Would you like a summary of your prediction accuracy? (Circle one answer.)
    Yes
    No

PLEASE DO NOT DISCUSS THIS STUDY WITH OTHERS. Don't forget to pick up your ten dollars on the way out. Thanks.

APPENDIX B

Input for Creation of Case Set II
(Cue List printed from terminal)

RATIFL2 DATA A1 06/25/84 9:53 ABER F 58

9.7	19.7	256.6	26.4	8.4	12.3
10.7	13.1	40.8	45.8	18.5	10.7
5.4	5.5	-47.9	9.7	-1.6	4.0
5.2	7.9	6.9	19.8	7.2	7.3
3.3	8.9	-7.5	28.3	18.3	13.5
6.6	7.6	-52.4	31.0	6.4	9.6
6.8	-2.0	-107.0	-46.6	-18.3	1.4
8.6	14.0	50.3	42.9	6.4	9.6
3.1	6.3	-53.1	-9.1	16.9	14.3
4.9	17.8	-15.6	-44.4	14.7	14.0
3.6	4.3	-50.9	6.5	-1.6	4.0
4.8	13.0	21.6	12.1	18.2	17.1
4.6	9.0	-42.9	-7.9	8.9	10.5
3.7	25.2	36.4	45.5	12.6	12.4
8.1	1.0	-82.7	8.1	-2.9	6.2
7.3	25.6	22.9	55.1	7.5	11.6
5.9	9.1	-52.2	-1.1	10.8	7.5
3.9	14.0	-19.1	19.1	9.3	12.3
5.3	13.2	-21.3	13.7	7.2	7.3
2.9	12.0	-12.6	26.1	7.5	11.6
1.8	19.9	71.3	26.6	17.6	14.9
6.1	17.6	7.1	24.8	7.5	11.6
3.9	3.7	-55.6	91.7	5.8	10.5
2.8	11.5	-26.5	5.3	-1.1	9.6
6.4	2.5	-81.3	-3.0	8.9	10.5
3.5	4.0	-48.1	72.9	17.6	14.9
4.7	8.0	-26.3	14.9	10.8	7.5
3.1	15.9	66.0	80.0	16.0	12.5
2.9	12.8	-19.0	31.7	17.6	14.9
1.8	20.2	224.7	-27.8	14.7	14.0

Example of a Case
(Printed directly from terminal)

```
CASE SET I            FORECASTING RETURN ON TOTAL ASSETS            FIRM NUMBER
                                                                         12

        CUES:                                                        RATIOS:

  1.    OPERATING EXPENSES/SALES.                                          7.1
  2.    WORKING CAPITAL PROVIDED BY OPERATIONS/SALES.                     10.1
  3.    OPERATING INCOME BEFORE DEPRECIATION/SALES.                       14.8
  4.    OPERATION INCOME/SALES.                                           11.8
  5.    NET INCOME/SALES.                                                  7.0
  6.    EFFECTIVE TAX RATE.                                               40.7

JUDGEMENT:  PLEASE ESTIMATE THE EXPECTED RETURN AT 12/31/X5.

CONFIDENCE:  PLEASE ENTER YOUR LEVEL OF CONFIDENCE.
             (1 2 3 4 5 6 7 : NO CONFIDENCE-EXTREME CONFIDENCE)

             PRESS ENTER TO CONTINUE.
```

Output from Cue Selection/Use Experiment
(Participant number 8)

HOLL2 DATA A0 06/28/84 20:21 ABER V 80 180 RECS VA TEC

1	BUTCH	HOLLAND	228-84-4554	27	3	1	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	5	3	2	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	8	4	3	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	8	3	4	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	11	4	5	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	9	3	6	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	18	4	7	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	10	2	8	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	2.5	4	9	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	20	5	10	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	22	5	11	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	21	5	12	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	19	2	13	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	16	5	14	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	29	4	15	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	20	4	16	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	14	4	17	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	8	4	18	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	19	5	19	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	16	5	20	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	13	4	21	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	4	4	22	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	12	4	23	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	10	2	24	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	9	3	25	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	15	4	26	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	4	5	27	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	1	5	28	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	10	5	29	4	9	18	19	20	26
1	BUTCH	HOLLAND	228-84-4554	15	5	30	4	9	18	19	20	26
2	BUTCH	HOLLAND	228-84-4554	18	3	1	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	14	4	2	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	5	5	3	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	7	4	4	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	8	4	5	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	5	5	6	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	-8	4	7	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	15	4	8	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	4	4	9	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	15	5	10	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	3.5	5	11	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	14	5	12	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	7	5	13	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	26	5	14	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	0	4	15	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	26	5	16	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	8	3	17	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	14	5	18	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	12	3	19	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	10	3	20	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	20	4	21	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	17	3	22	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	4	5	23	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	9	4	24	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	1.5	3	25	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	3	4	26	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	7	3	27	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	17	5	28	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	13	3	29	1	2	3	4	5	6
2	BUTCH	HOLLAND	228-84-4554	19	4	30	1	2	3	4	5	6

Correlation Statistics
(Participant 8, Case Set I)

DEP VARIABLE: YE

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB>F
MODEL	6	546.991	91.165202	5.528	0.0011
ERROR	23	379.298	16.491237		
C TOTAL	29	926.290			

ROOT MSE	4.060940	R-SQUARE	0.5905
DEP MEAN	13.763333	ADJ R-SQ	0.4837
C.V.	29.5055		

VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR H0: PARAMETER=0	PROB > |T|
INTERCEP	1	-0.986908	3.709863	-0.266	0.7926
X4	1	0.882476	0.299250	2.949	0.0072
X9	1	3.110555	1.312864	2.369	0.0266
X18	1	0.016572	0.011110	1.492	0.1494
X19	1	-0.00249047	0.086300	-0.029	0.9772
X20	1	-0.00586697	0.071390	-0.082	0.9352
X26	1	0.188939	0.132272	1.428	0.1666

DEP VARIABLE: YS

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB>F
MODEL	6	1195.003	199.167	18.910	0.0001
ERROR	23	242.238	10.532101		
C TOTAL	29	1437.242			

ROOT MSE	3.245320	R-SQUARE	0.8315
DEP MEAN	13.183333	ADJ R-SQ	0.7875
C.V.	24.61684		

VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR H0: PARAMETER=0	PROB > |T|
INTERCEP	1	-2.408927	2.964755	-0.813	0.4248
X4	1	1.129501	0.239147	4.723	0.0001
X9	1	1.741023	1.049182	1.659	0.1106
X18	1	0.018632	0.008878361	2.099	0.0470
X19	1	-0.036831	0.068967	-0.534	0.5984
X20	1	0.068561	0.057052	1.202	0.2417
X26	1	0.173066	0.105706	1.637	0.1152

Correlation Statistics
(Continued)

CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0

	YE	YS
YE	1.00000 0.0000	0.81482 0.0001
YS	0.81482 0.0001	1.00000 0.0000

CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0

	YEHAT	YSHAT
YEHAT PREDICTED VALUE	1.00000 0.0000	0.95694 0.0001
YSHAT PREDICTED VALUE	0.95694 0.0001	1.00000 0.0000

CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0

	YE	YSHAT
YE	1.00000 0.0000	0.73536 0.0001
YSHAT PREDICTED VALUE	0.73536 0.0001	1.00000 0.0000

The vita has been removed from the scanned document