FACTORS AFFECTING THE ACCURACY

OF EFFICACY JUDGMENTS

By

GERALD L. HARBOUR

Bachelor of Arts
Western Washington College
Bellingham, Washington
1972

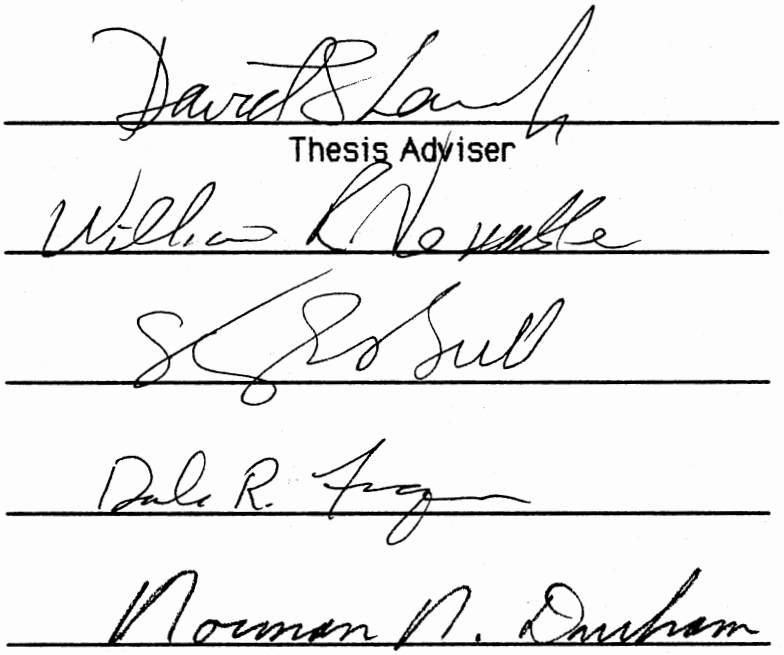
Master of Science
Eastern Washington College
Cheney, Washington
1978

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 1988

FACTORS AFFECTING THE ACCURACY

OF EFFICACY JUDGMENTS

Thesis Approved:



Thesis Adviser

Dean of the Graduate College

ACKNOWLEDGMENTS

Rarely does the attainment of a difficult goal represent only an individual effort. Rather, such goals are achieved through the collective effort, knowledge, and support of many individuals. Acknowledging this truth, I wish to thank several individuals who contributed in many different ways to my doctoral program at Oklahoma State University.

I would first like to express my sincere appreciation to Dr. David Lane and Dr. William Venable for their encouragement, advice, and friendship over these past few years. I also wish to thank Dr. Kay Bull for his assistance, interest, and kindness shown me during my graduate program. Many thanks are also extended to Dr. Nancy Fagley and Dr. Dale Fuqua for their helpful suggestions and critical comments concerning the dissertation.

I wish to also thank the general faculty of the Department of Applied Behavioral Studies. Their expertise, professionalism, and warmth as human beings will always be remembered and appreciated.

Finally, I would like to thank family, friends, and special ones for their emotional and physical support during my graduate program. These expressions of thoughtfulness and caring meant more than words will ever be able to express.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Chapter I provides a general overview of the study. The chapter is divided into seven sections. Section I begins by briefly exploring the construct of self-efficacy, the topic of this study. Following this introduction, subsequent sections describe the study purpose and problem statement (Section two), a short description of the study (Section three), definitions of key terms (Section four), limitations and assumptions underlying the study (Section five), hypotheses that will be tested (Section six), and the significance of the study (Section seven). Section eight concludes the chapter by providing an overview of the organization of the study.

Section One: Self-Efficacy

A fairly accurate assessment of one's own performance potential is of considerable value for successful functioning. As noted by Bandura (1986), large misjudgments of performance potential in either direction can have potentially serious consequences. People who grossly overestimate their performance capabilities for example, may undertake activities or expect performance attainments that are clearly beyond their reach. Failure to attain such performance levels may result in disappointment, heightened levels of frustration, or

even task abandonment. Conversely, people who greatly underestimate their performance capabilities commonly display self-limiting behaviors, including learned helplessness.

How people judge their own performance capabilities has often been described under the construct of self-efficacy. Self-efficacy is a key concept in Bandura's (1977) social learning theory and refers to personal judgments of performance capabilities in a given domain of activity (Bandura, 1977, 1981, 1982). According to Bandura (1986), self-efficacy is not concerned with the skills one possesses but rather with the personal judgment of what one can do with those skills. Competent performance requires both the effective use of skills and self-beliefs of efficacy.

Numerous studies have documented the ability of self-efficacy judgments to predict subsequent performance outcomes (see general review by Gist, 1987). Although research has demonstrated the reliability of such predictions, it has yet to focus specifically on how efficacy predictions are actually made or what factors may affect their accuracy. This lack of research has previously been discussed by Bandura (1986) and Schunk (1985) who both suggested that there is little understanding of how people process multidimensional, efficacy information. According to the two authors, research is needed to determine how people select, weight, and combine efficacy information from diverse sources. Specifically, they suggested initial research is needed to answer two important questions: (1) Do people choose common themes when selecting and weighting efficacy information or are there distinct, individual differences? and (2) Do

individuals with poor efficacy judgments weight and select efficacy information differently than individuals with more realistic and accurate efficacy judgments?

Although Bandura (1986) and Schunk (1985) focused their attention on the importance of efficacy information cue selection and weighting, other factors may also affect the accuracy of self-predictions of performance. Studies by Ward and Eisler (1987) for example, demonstrated that individuals identified as possessing a Type A personality behavior frequently overestimated their own performance capabilities, setting unrealistic and unobtainable goals for themselves. In contrast, Type AB and Type B individuals frequently underestimated or accurately predicted their own performance capabilities. Ward and Eisler concluded from their studies that additional research is needed to explore the links between individual differences such as the Type A - Type B behavior continuum, and the construct of self-efficacy.

Section Two: Study Purpose and Problem Statement

Following the suggestions of Bandura (1986), Schunk (1985), and Ward and Eisler (1987) concerning additional research into factors affecting the accuracy of efficacy judgments, the following study was developed. The overall objective of the study was to begin to understand how individuals make predictions of personal performance outcomes and what specific individual differences may affect the accuracy of these predictions. The specific purpose of the study was to investigate how college students majoring in education

make self-predictions of their own test performance and how gender, age, year in school, perceived sex-role identity, test anxiety level, Type A - Type B behavior pattern, grade point average, and actual test performance affect the accuracy of these predictions. The study attempted to determine which information sources were selected during the prediction process and how they were weighted. The study further attempted to identify specific variables affecting the accuracy of efficacy judgments.

The dependent variable in the study was the accuracy of predicted test performance, defined as the difference between the student's actual test score and predicted test score. The independent variables were the types of information sources reported by students in making their test predictions, how such information sources were weighted, age, gender, grade point average, year in school, perceived sex-role identity, test anxiety level, Type A - Type B behavior pattern, and actual test performance. Possible sources of efficacy information that might have been selected and weighted by students included:

- teacher effectiveness in presenting course material
- how well other students appear to be doing in the course
- perceived test difficulty
- self-confidence level of the student
- concentration level while studying
- general academic performance
- anxiety level
- knowledge of the material to be tested

- past performance in other educational courses
- the interest of the tested material to the student
- the mood of the student
- the amount of encouragement given to the student by the instructor
- the student's physical health
- the perceived effectiveness of the student's study skills
- how lucky the student feels
- the amount of effort exerted by the student in studying for the test
- how well the student has done on previous tests in the course.

The underlying question of the study asked whether there were differences in the types of information sources selected and/or weighted among students who made accurate predictions of test scores when compared to those who did not. A second research question asked whether age, gender, year in school, perceived sex-role identity, test anxiety level, Type A - Type B behavior pattern, grade point average, and actual performance attainments could be correlated with the accuracy of predicted test performance.

Section Three: Study Overview

Immediately before each of four, regularly scheduled, class examinations, subjects were provided with a questionnaire that asked them to predict their own test performance and to state how confident they were in their prediction. They were also asked to rate how influential each of seventeen information sources were in determining their prediction. Finally, they were asked to identify and rank the four most critical information sources used in making their

test prediction. Also during the semester, but not on any of the regularly scheduled examination days, all subjects were given (1) the Bem Sex-Role Inventory (Bem, 1981), a measure of perceived sexual role identity; (2) the Test Anxiety Inventory (Speilberger, 1980), a measure of trait anxiety concerning test taking; and (3) the Jenkins Activity Survey, Form T (Krantz, Glass, and Snyder, 1974), a measure of the Type A - Type B personality behavior continuum for university students.

Section Four: Definition of Terms

For the purpose of this study, the following definitions were used:

Self-efficacy: Self-efficacy, as defined by Bandura (1982), refers to personal judgments of performance capabilities in a given domain of activity. It is a personal prediction of how well one can perform actions in specific situations.

Predicted test score: A predicted test score is a self-reported estimation of how well a student thinks she or he will do on a 40-point multiple-choice test. Only one number from 0 to 40 may be selected in deriving a predicted test score.

Actual test score: An actual test score is the raw score achieved on a 40-point, multiple-choice test. Raw scores may range from 0 to 40.

Accuracy score: An accuracy score equals the actual test score minus the predicted test score. Accuracy scores may be either positive or negative.

Cumulative accuracy score: The cumulative accuracy score equals the

sum of accuracy scores obtained over four separate tests.

Confidence of prediction: The confidence of prediction is a self-reported estimation of the confidence of an individual in obtaining a predicted test score. The confidence of prediction may range from 0 (a complete lack of confidence) to 100% (complete confidence).

Information cue: An information cue is information that is selected, weighted, integrated, and transformed along with other information cues into a predicted test score. For the purposes of this study, information cues refer only to the seventeen information sources identified on the Exam Prediction Questionnaire.

Critical information cue: A critical information cue is an information cue designated by subjects as being especially important in formulating a predicted performance score. Of the seventeen information cues, subjects may designate only four as critical information cues.

Masculinity: Masculinity is a behavioral trait associated with an instrumental orientation and a cognitive focus on getting the job done or the problem solved (Parsons and Bales, 1955). Masculinity is not gender specific.

Femininity: Femininity is a behavioral trait associated with an expressive orientation, an affective concern for the welfare of others, and the harmony of the group (Parsons and Bales, 1955). Femininity is not gender specific.

Androgyny: Androgyny is a behavioral trait associated with being both compassionate and assertive, expressive and instrumental, and

feminine and masculine depending upon the situational appropriateness of these modalities (Bem, 1974).

Type A behavior pattern: For the purposes of this study, a Type A behavior pattern is operationally defined as a score of greater than 10 on the Jenkins Activity Survey, Form T. A similar identification process was used by Krantz, Glass, and Schaeffer (1974) and Ward and Eisler (1987).

Type AB behavior pattern: For the purposes of this study, a Type AB behavior pattern is operationally defined as a score of between 6 and 10 on the Jenkins Activity Survey, Form T. A similar identification process was used by Ward and Eisler (1987).

Type B behavior pattern: For the purposes of this study, a Type B behavior pattern is operationally defined as a score of less than 6 on the Jenkins Activity Survey, Form T. A similar identification process was used by Krantz, Glass, and Schaeffer (1974) and Ward and Eisler (1987).

Section Five: Limitations and Assumptions

The following limitations in the study must be considered:

1. The research design is correlational. As noted by Issac and Michael (1981), correlation does not necessarily imply causation.
2. The study is limited to the population under investigation. All subjects are education majors at a large, mid-western, public university. It cannot be assumed that the studied group are representative of all university students or the general population.
3. Subjects were not randomly selected. Instead, intact groups were

used to facilitate the logistics of a multiple measure study. However, the intact groups are highly representative of the population described in limitation #2.

4. There is a greater proportion of females (71%) than males (29%) in the studied group. This ratio, however, is typical among the studied population.

5. The possible range of information cues selected by subjects was predetermined. Open ended self-reports were not used. However, as described in Chapter III, considerable care was taken to provide study participants with as wide a selection of efficacy information cues as possible.

Certain assumptions were made during the study.

1. Self-reports by subjects are accurate reflections of their actual opinions. Steps were taken during the study to decrease the likelihood of self-serving distortions. For example, subject identity was protected at all times.

2. Subjects, as demonstrated during pilot testing and explained in Chapter III, understood and comprehended the instruments used in the study.

3. There is a linear relationship between predicted test score and actual test score. The validity of this assumption was confirmed via scatterplots during the statistical analysis phase of the study.

Section Six: Hypotheses

The following hypotheses were tested. Null hypotheses and acceptable alpha levels are formally stated in Chapter IV.

1. There is no correlation between predicted test scores and actual test scores.
2. There is no correlation between confidence of predictions and accuracy scores.
3. There is no correlation between predicted test scores and confidence of predictions.
4. There is no correlation between the independent variables of gender, age, year in school, perceived sex-role identity, year in school, test anxiety level, Type A - Type B behavior pattern, grade point average, and actual test performance and the dependent variable cumulative accuracy scores.
5. There is no correlation between information cue rating and accuracy scores during each of the four examinations.
6. There is no correlation between critical information cues selected, and accuracy scores during each of the four examinations.
7. There are no significant differences among the ratings of the seventeen information cues.

Section Seven: Study Significance

As noted earlier, possessing a reasonably accurate appraisal of one's own performance capabilities is of considerable value for successful functioning. Major misjudgments in either direction can have serious consequences. For example, a jet fighter pilot who flies beyond his own capabilities can quickly find himself in a dangerous, if not potentially fatal, situation. Business people who oversell their ideas and abilities can not only damage their own careers but can also

seriously impact the corporations they work for.

Understanding the factors affecting the accuracy of efficacy judgments could improve decision making when self-predictions of ability must be considered. A better understanding of how efficacy judgments are made may also allow professionals to more effectively assist individuals who either grossly overestimate or underestimate their own performance capabilities.

Section Eight: Organization of the Study

This chapter briefly introduced the subject of self-efficacy, the focus of the study. It further described the purpose of the study and the problem statement, gave a brief overview of the study, noted the limitations and assumptions made concerning the study, presented the seven hypotheses to be tested, and briefly explored the significance of the study.

Chapter II contains a review of the current literature on self-efficacy judgments, focusing especially on the information sources used in making self predictions. The contributions of attribution theory to this problem are also explored. The potential impact of gender and sex-role identity on performance predictions is also considered. Finally, Chapter II examines how the Type A - Type B personality behavior continuum may affect the accuracy of efficacy judgments.

Chapter III includes a description of the subjects, the instrumentation used, procedures followed during the study, and a discussion of the research design. Chapter IV presents the results of

the study, including all statistical analyses of the collected data. Chapter V discusses and interprets the statistical results and offers possible explanations for the observed phenomena. It also provides suggestions for future research concerning the identification of factors affecting the accuracy of efficacy judgments.

CHAPTER II

REVIEW OF LITERATURE

The review of literature presented in Chapter II is divided into six sections. The chapter begins by providing a definition and overview of self-efficacy. The literature citing the use of self-efficacy judgments as predictors of task performance is reviewed in Section two. Following these two sections, Section three focuses on identifying potential information sources that individuals may select in making efficacy judgments. The review will begin with the work of Bandura (1982) on self-efficacy cues. Heeding Schunk's (1985) advice that findings from attribution theory may also be relevant, an overview of the literature concerning attribution theory will be presented and specific findings of perceived causal factors of success in the achievement domain summarized. Section four explores the role that gender and sexual-role identity may play in affecting self-efficacy percepts. Section five describes the influence of the Type A - Type B behavior pattern on the accuracy of efficacy judgments. Finally, Section six summarizes the reviewed materials.

Section One: An Overview of Self-Efficacy

As noted by Bandura (1986), psychological theories and research commonly focus on issues concerning either the acquisition of knowledge or the performance of response patterns. Little effort,

however, has been directed toward understanding the process governing the interrelationship between knowledge and action. In an attempt to bridge this gap, some researchers have concentrated their attention on the mediating role that an individual's own thoughts play between knowledge and action. According to Bandura (1986), the issues addressed in this line of inquiry are all concerned with, " how people judge their capabilities and how their self-percepts of efficacy affect their motivation and behavior" (p. 391). Efforts along this line of inquiry include the work of DeCharms (1978), Garber and Seligman (1980), Lefcourt (1976), Perimuter and Monty (1979), and Rotter, Chance, and Phares (1972).

As noted in Chapter I, self-efficacy is a key concept in Bandura's (1977) social learning theory. As previously defined, self-efficacy is a personal judgment of how well one can perform actions in specific situations (Bandura, 1982). Self-efficacy does not concern itself with the skills one possesses but rather with the personal judgments one makes concerning the application of those skills (Bandura, 1986).

According to Bandura (1982), self-efficacy affects task choice, task effort, and task persistence. First, regarding the effects of self-efficacy judgments in determining task choice, various studies in the area of career choice including those by Betz and Hackett (1981), Mitchell and Krumboltz (1984), Molnar and DeLauretis (1973), and Wheeler (1983), have identified self-efficacy as a critical factor in the perceived range of career options. These studies have shown, among other things, that low perceived self-efficacy of mathematical

ability can keep many prospective students from entering scientific or engineering fields.

As previously mentioned, judgments of self-efficacy can also determine the amount of effort people will expend on a task and how persistent they will be in the face of obstacles or aversive experiences. High self-efficacy judgments normally coincide with greater task effort and expenditure. Findings by Bandura and Cervone (1983), Brown and Inouye (1978), Schunk (1984) and Weinberg, Gould, and Jackson (1979) indicate that people with low self-efficacies exert less effort or give up altogether when confronted with task obstacles. Lent and Larkin (1984) discovered, for example, that subjects reporting high self-efficacy for educational requirements generally achieved higher grades and persisted longer in technical/scientific majors over the following year than those with low self-efficacy. High task effort and persistence apparently contributed to higher performance attainments.

Bandura (1982) suggested that self-efficacy judgments can vary on three important dimensions. These dimensions include magnitude, strength, and generality. Magnitude refers to the level that people believe they can attain. It represents a prediction of a personal performance outcome. Thus, some people may have high self-efficacies or expectations for certain tasks, whereas other people may have lower expectations. Strength concerns whether the conviction regarding magnitude is strong or weak. It is commonly stated as a confidence level from 0% to 100%. Weak percepts of self-efficacy, according to Bandura (1986), can be easily extinguished

by unpleasant experiences, whereas people possessing strong percepts of self-efficacy will persevere despite mounting difficulties. Generality indicates the degree to which expectations are generalized across situations. Some individuals may judge themselves efficacious only in specific domains. Others may judge themselves efficacious across a wide variety of situations and events. Self-efficacy is normally measured as it relates to specific tasks. Consequently, in most studies only magnitude and strength determinations are made. Little research has been conducted on how efficacy perceptions can be generalized (Gist, 1987).

Self-efficacy measurements are normally taken immediately before a specific task (Bandura, 1982). Subjects are asked to estimate their expected level of performance, as well as the strength or confidence level of their estimation. In some instances, subjects are asked for strength of magnitude for various levels of performance. In other instances, subjects select their expected attainment level and report a confidence level for this estimation.

Self-efficacy is an important construct in Bandura's (1982) social learning theory. Self-efficacy may be defined as an individual's perception of their probability of success at a subsequent task. Self-efficacy affects task choice, task effort, and task persistence. Self-efficacy judgments may be measured along three dimensions: level, strength, and generality.

Section Two: Predicting Task Performance

Bandura (1982) indicated that self-efficacy can predict

performance in a variety of domains if the efficacy measure is tailored to the specific tasks being measured. Numerous studies support this conclusion, reporting significant correlations between self-efficacy and subsequent task performance (Bandura, 1982; Bandura , Adams and Beyer, 1977; Bandura, Adams, Hardy, and Howells, 1980; Barling and Beattie, 1983; Chambliss and Murray, 1979; Covington and Omelich, 1979; Feltz, 1982; Gould, Weiss and Weinberg, 1981; Locke, Frederick, Lee, and Boko, 1984; Schunk, 1984; Schunk and Gunn, 1986; Siegel, Galassi, and Ware, 1985: and Taylor, 1987).Research into the use of self-efficacy as a performance indicator has been primarily confined to three major areas: sports, business, and academic settings. Studies in each of these three areas are briefly summarized.

First, in regards to sports research, self-efficacy has been demonstrated to be a reliable predictor of subsequent performance. Nelson and Furst (1972) for example, in a study utilizing a test of arm strength, found that weaker male subjects who believed, along with their opponents, that they were stronger outperformed their opponents 83% of the time. Mahoney and Avener (1977) reported that national caliber gymnasts who were uncertain of their ability did less well than more efficacious gymnasts.

Studies conducted by Weinberg, Gould, and Jackson (1979), Weinberg, Yukelson, and Jackson (1980), and Weinberg, Gould, Yukelson, and Jackson (1981), found a causal relationship between performance in a motor task and self-efficacy. The research involved subjects who competed against a confederate in a measure of leg

strength. Self-efficacy was manipulated by the confederates who stated that they were either varsity track athletes or had injured legs. Subjects who thought they were competing against varsity athletes reported low self-efficacies whereas subjects who thought they were competing against injured opponents reported high self-efficacies. Results in all three studies indicated that self-efficacy was a good predictor of task effort. High self-efficacy subjects performed the leg strength task significantly longer than subjects with low self-efficacy.

Taylor (1987) investigated the use of self-efficacy, along with state and trait anxiety, to predict performance among varsity athletes at the University of Colorado. Results indicated that self-efficacy was a significant predictor of performance in a variety of sports.

Highlen and Bennett (1979) and Gould, Weiss, and Weinberg (1981) in separate studies, both found self-efficacy to be a significant predictor of athletic performance among Big Ten wrestlers. Wrestlers finishing in the top three places during Big Ten tournament action commonly indicated higher efficacy judgments preceding the tournament than did wrestlers who finished lower.

Research in sports psychology seems to support the idea that self-efficacy can be an accurate predictor of subsequent sports performance. Individuals indicating high self-efficacy percepts frequently realize those expectations in actual competition. Conversely, athletes expressing doubts about their own abilities commonly finish lower in competitive standings despite the fact that

they may actually be the superior athlete.

The second area where self-efficacy has been studied is business. Research into the predictive ability of self-efficacy in this setting, however, has not been nearly as extensive as in the sports domain (Gist, 1987). One notable exception to this paucity is the study of Barling and Beattie (1983). They found that efficacy judgments of life insurance sales representatives were significantly correlated with the number of calls made per week, the number of policies sold, sales revenue, and a composite performance index.

In a somewhat related study of faculty research productivity at a large, eastern, public university, Taylor, Locke, Lee, and Gist (1984) found self-efficacy to be a reliable predictor of the number of publications published per year by faculty members. Once again, self-efficacy judgments fairly accurately predicted subsequent performance.

In the third area, academic settings, numerous studies have been conducted in the use of self-efficacy judgments as a predictor of academic performance. Schunk (1984), in a study of mathematics ability and self-efficacy among children, found children's perceptions of self-efficacy to have a positive relationship to subsequent skilled performance. More rapid problem solving during training and task performance were associated with higher self-efficacy.

Lent, Brown, and Larkin (1986) found that self-efficacy contributed significantly to the prediction of technical grades, academic persistence, and range of career options considered among college undergraduates majoring in science and engineering. In a

hierarchial regression analyses used by the authors, self-efficacy remained significant even after math ability, high school achievement, and vocational interest had been removed from the regression equation.

Siegel, Galassi, and Ware (1985) conducted a study contrasting the ability of two theoretical models to explain mathematics final examination performance. The authors hypothesized that Bandura's (1977) social learning model would account for more variation in math performance than the math apptitude-anxiety model. Results from the study confirmed the stated hypothesis, indicating that significantly more variance was explained by the social learning variables than the math aptitude-anxiety variables. Self-efficacy accounted for a significant amount of the variance in the social learning model.

In other related research, Locke, Frederick, Lee, and Bobko (1984) conducted studies exploring the effects of self-efficacy, goals, and task strategies on goal choice and task performance using college undergraduates as subjects. The study found that self-efficacy was a significant predictor of future performance. Specific findings indicated that self-efficacy ratings for moderate to difficult levels of performance were the best predictors of future performance.

Finally, Covington and Omelich (1979) found that adults' personal expectations of successful performance was one of the best predictors of later performance. Perceptions of self-efficacy were shown by Covington and Omelich to be a more reliable indicator of

subsequent performance than statements of causal attributions.

To summarize this section, the literature from sports psychology, business, and academic settings supports Bandura's (1982) contention that self-efficacy judgments can be a reliable predictor of performance in a variety of domains if the efficacy measure is adapted to the specific tasks being measured. This latter condition seems especially critical for the successful prediction of subsequent performance. Numerous studies have found significant correlations between self-efficacy percepts and subsequent task performance. An individual's own judgments of potential performance outcomes can frequently be an accurate predictor of actual performance outcomes.

None of the reviewed studies, however, investigated specifically how self-efficacy judgments are made or what information sources are selected, weighted, and integrated by an individual when determining an efficacy judgment. Nor have these studies attempted to determine why some individuals can fairly accurately predict their own performance outcomes, whereas other individuals persistently under-rate or over-rate their own abilities. A major line of inquiry to be explored in section three of this chapter is assessing how individuals may develop judgments of self-efficacy.

Section Three: Formulating Efficacy Judgments

Two major sources of potential information concerning the development of efficacy judgments will be reviewed in section three. The first source of reviewed literature summarizes mainly the work

of Bandura (1986) on self-efficacy judgments. The second major source of reviewed literature deals with attribution theory. Although causal attributes occur post-performance whereas self-efficacy judgments are pre-performance, such attributes may affect subsequent efficacy judgments.

Bandura's Concepts

Bandura (1982) identified four principle information sources that influence self-efficacy judgments. In decreasing order of importance these include: (1) enactive mastery, defined as repeated performance accomplishments; (2) vicarious experiences or observing others; (3) verbal persuasion; and (4) emotional or physiological arousal. Such efficacy cues, however, are only instructive once they have been cognitively appraised (Bandura, 1986). Bandura cautioned that a distinction must be made between information conveyed by environmental events and information that is selected, weighted, and integrated into self-efficacy judgments.

Enactive Master. Findings by Bandura, Adams, and Beyer (1977), Biran and Wilson (1981), and Feltz, Landers, and Raeder (1979) sugggest that enactive accomplishments provide the most influential influential source of efficacy information. The reason for this, as suggested by Bandura (1986), is that enactive attainments are based on authentic mastery experiences. Enactive mastery has been shown to enhance self-efficacy more than any of the other efficacy cues. Past successes normally raise efficacy expectations, while repeated failures lower them.

According to Bandura (1982), mastery is facilitated when gradual accomplishments build the skills and coping abilities needed for task performance. As noted by Gist (1987), however, although successful performance can be a powerful enhancer of self-efficacy, in some situations, individuals may not expose themselves to opportunities for enactive mastery. This observed hesitancy may be caused by fear or some personal incapacity.

Bandura (1986) points out that perceived efficacy is not only affected by past successes and failures but also by biases in the monitoring of the experiences themselves. Individuals who selectively remember only their past successes will frequently overestimate their self-efficacy judgments. Conversely, people who selectively remember only their failures will frequently underestimate estimate their efficacy judgments.

Vicarious Experiences. A second, although somewhat less influential source of efficacy information, comes from vicarious experiences. Bandura, Adams, Hardy, and Howells (1980) and Kazdin (1979) have demonstrated that people who see or mentally visualize other similar people perform successfully can raise their own self-percepts. Apparently, by watching similar others, people can persuade themselves that if others can do it, so can they. These same studies also suggest that modeling is more effective when the models succeed after overcoming initial difficulties than when they exhibit exemplary performance from the beginning.

Bandura (1977) also notes that the effects of vicarious experiences are enhanced when the modeled behavior produces clear

results or consequences. Further, self-efficacy is increased when there is a similarity between the model and the subject in terms of age, capabilities, and other personal characteristics. Like enactive experiences, modeled successes by similar others generally raises self-efficacy judgments, whereas modeled failures lower self-appraisals of efficacy.

Bandura (1986) listed two conditions under which self-efficacy appraisals are especially sensitive to vicarious information. The first condition involves the amount of uncertainty one possesses about one's own capabilities. It appears that perceived self-efficacy can be easily changed through modeling when individuals have had little or no experience on which to base evaluations of their personal competence.

The second condition involves the criteria by which ability is evaluated (Festinger, 1954; Suls and Miller, 1977). Bandura (1986) notes that when factual evidence for acceptable performance is lacking, personal efficacy must be calibrated in terms of the performance of others. Since most performances are evaluated in social terms (how well one person does in comparison to another), social comparison information is an important self-efficacy cue. Vicarious experiences appear to be an important efficacy cue. Observing another person's task performance can frequently raise or lower our own performance expectations. This conclusion appears especially true if the observed individual is similar in many ways to ourselves.

Verbal Persuasion. The third source of efficacy information

according to Bandura (1982) is verbal persuasion. Verbal persuasion is aimed at convincing a person of his or her capability of performing a task. Bandura (1986) cautions that verbal persuasion alone may be of limited power in increasing perceptions of self-efficacy. Verbal persuasion can, however, contribute to successful performance if the heightened appraisal is within realistic boundaries.

Bandura (1986) notes that it is probably more difficult to produce lasting increases in perceived efficacy by verbal persuasion than it is to undermine it. Apparently, individuals who have been persuaded of their inefficacy tend to avoid challenging activities as well as to give up easily in the face of difficulties.

Physiological State. Finally, Bandura (1982) suggests that individuals' perceptions of their physiological state may be used in assessing performance capability. An individual in an aroused state, such as someone experiencing high visceral anxiety while taking a test, may interpret the arousal as debilitating fear and feel excessively vulnerable to failure (Gist, 1987). Bandura (1986) suggests that because high arousal usually debilitates performance, people are more inclined to expect failure if they are tense and viscerally agitated. Conversely, people are more inclined to expect success when they are not adversely affected by aversive arousal.

Bandura further notes that several factors affect the cognitive processing of physiological efficacy information. These include the sources of arousal, the level of activation, the circumstances under which arousal is elicited, and past experiences of how arousal affected one's performance.

Work by Hollandsworth, Glazeski, Kirkland, Jones, and van Norman (1979), suggests arousal cues are frequently interpreted by their perceived effect on performance. For those people who generally find arousal beneficial, arousal will have a different efficacy meaning than for those who view arousal as adversely affecting their performance. High achievers apparently view arousal as a facilitator, whereas low achievers view it as a debilitator.

Perceived physiological state, as suggested by Bandura (1986), may represent a poor source of efficacy information. Individuals who focus on their own anxiety prior to a task may significantly underestimate their own abilities and expected performance attainments. Such a focus may also affect original task choice since tasks associated with anxiety producing feelings may be purposefully avoided.

To summarize Bandura's (1982) ideas, there are four principle sources of information that people select, weight, and integrate when making self-efficacy judgments. In decreasing order of importance, these efficacy sources or cues are enactive experiences, vicarious experiences, verbal persuasion, and physiological state. In all instances, positive experiences, either enactive or vicarious, appear to increase efficacy judgments, whereas negative experiences seem to lower efficacy perceptions. Also, focusing on one's physiological or emotional state may lead to an underestimation of ability.

Contributions of Attribution Theory

As noted by Bandura (1982), past performances are a valuable

source of efficacy information. The important point of such experiences is, however, not what actually took place but rather how they are perceived and interpreted to have taken place. Identifying causal factors of past experiences may be incorporated into future efficacy judgments. This concept is suggested by Schunk (1984), who hypothesizes that an individual's attributions concerning past performance or failure may influence subsequent self-efficacy judgments.

Attribution theories of behavior (Heider, 1958; Kelley, 1967) suggest that individuals make causal ascriptions for the outcomes of their actions. Future performance expectancies such as self-efficacy, may depend heavily on causal ascriptions (Weiner, 1979). Schunk (1985) illustrates this concept by stating, "if one believes that the task circumstances will remain much the same, attributing prior successes to relatively stable causes such as high ability or low task difficulty should result in higher expectations of future success than attributions to the more unstable causes of great effort or good luck (McMahn, 1973; Weiner, Nierenberg, and Goldstein, 1976)" (p. 212).

According to Weiner (1985), humans feel compelled to attribute past performance outcomes to some causal factor(s). One possible explanation for this phenomena, as suggested by White (1959), is that individuals need to understand themselves and their surrounding environment. White termed this motivation the principle of mastery.

Also, as pointed out by Weiner (1985), it is functional to know why an event has occurred. As stated by Kelly (1971): "The attributor is not simply an attributor, a seeker after knowledge; his latent goal

in attaining knowledge is that of effective management of himself and his environment" (p. 22). Once a cause or causes of a performance outcome have been determined, effective self-management may become possible and a guide for future action generated (Weiner, 1985).

According to Weiner (1985), if the prior performance outcome was a success, there is a natural tendency to reinstate the prior causal network. Conversely, if the prior outcome was undesirable, there is a strong possibility that there will be an attempt to alter the causes to produce a different, more positive effect.

Many investigations (Weiner, 1985) have been conducted to systematically examine causal perceptions, particularly those involving the perceived causes and failures concerning achievement related situations. Many of these studies have focused specifically on the academic setting. As outlined by Weiner (1985), two research procedures have been used in these studies. In one approach, subjects are provided only with performance outcome information, namely that success or failure has occurred. The outcome chosen by the investigator may be imagined, induced, or have occurred in a real setting. The described outcome may pertain to the subject or to another who is being judged. Subjects are then asked to explain the outcome, using a free-response procedure where the possibilities that come to mind are listed.

In a second but somewhat different approach than the first, subjects are provided with a large list of causes and rate the contribution of each cause to the outcome. According to Weiner

(1985), these causes are often determined during pilot studies using a free-response methodology. The causes attempt to represent the dominant perceptions held by individuals. Eight of these investigations especially relevant to this study are briefly summarized and the attributions made by subjects noted. Frieze (1976) used college students as subjects and had them explain attributions for success and failure on hypothetical school and game performance. Performance attributions were obtained from both a self and other perspective. The dominant attributes used to explain task performance were effort, ability, luck, and the influence of other persons.

In a study by Elig and Frieze (1979), college students were used to solve a series of anagrams. Upon completing the task, they were then asked to explain their performance outcomes. Dominant attributions in this experiment included perceived task difficulty, ability, effort and mood. In yet another study designed by Frieze (Frieze and Snyder, 1980), first through fifth graders were asked to explain performance outcomes achieved by other students. The tasks involved a hypothetical academic test, an art project, sports, and a game. Dominant attributions reported included effort, ability, interest, and perceived task difficulty.

Cooper and Burger (1980) asked various teachers to identify the causal attributes affecting the academic performance of their students. The dominant attributes listed included typical and immediate effort, academic ability, and attention span. Burger, Cooper, and Good (1982) conducted a similar study in an attempt to

replicate their earlier findings. Once again, the teachers attributed student performance to immediate and stable effort, ability, and attention.

Anderson (1983) presented college students with a variety of hypothetical achievement situations and asked them to attribute causal factors to the observed performance outcomes. Dominant attributions included effort level, general knowledge, behavioral preparation, experience, and skill.

Wilson and Palmer (1983) conducted two studies involving college students. The students were asked to explain results of school examinations that they had taken. In the first study, dominant attributions identified included effort, luck or chance, task characteristics, and interest. The second study found similar findings, including effort, ability, task characteristics, and interest.

Bar-Tal, Goldberg, and Knaani (1984) investigated self-perceptions held by seventh graders concerning performance outcomes on academic tests. Two studies were conducted by the authors. In the first study, advantaged students were used. Their causal attributions for test performance included test preparation, amount of effort exerted in studying, their concentration level while studying, and the teacher's ability. In the second study, disadvantaged students were used. Their causal attributions are very similar to the first group, including test preparation, concentration level while studying, effort exerted during studying, and self-confidence level.

From the above investigations, a common theme seems to emerge. Although the research investigations made use of a wide

variety of types of subjects judging a variety of achievement situations, and involving the perspective of the self or other, causal attributions are remarkably similar. As stated by Weiner (1985), "a virtually infinite number of causal ascriptions are available in memory. However, within the achievement domain, a relatively small number from the vast array tend to be salient. The most dominant of these causes are ability and effort. That is, success is ascribed to high ability and hard work, and failure is attributed to low ability and the absence of trying" (p. 549). Triandis (1972) notes that these same attributions have been found in a number of different cultures.

Porac (1981), in a departure from most other attributional research, explored the intercausal relationships of performance attributes. Porac conducted two studies to determine whether students perceive meaningful influence patterns among the causal variables involved in explaining test performance. Students were requested in the study to specify the extent to which each of four performance causes (ability, effort, difficulty, and chance) affected the others. Three types of intercausal effects were presented by Porac: a negative relationship (e.g., ability is perceived to have reduced effort), a positive relationship (e.g., task difficulty is perceived to have increased effort), and no relationship (e.g., ability had no effect on luck).

In the first study, undergraduate college students were asked retrospectively to account for their midterm exam performance. Porac found from this initial study that students perceived a number of both unidirectional and bidirectional intercausal effects, and that

these were related to both perceived success and causal attributions. A second, similar study was also conducted. In the second study, however, subjects were requested to explain the exam performance of a hypothetical student. Results corroborating those of the first study were obtained.

Porac summarized his findings by making three general points. First, students clearly interpreted the effects of ability, effort, difficulty, and chance as an "interlocked set." Second, the specific types of intercausal effects take on significance in light of the relationship between perceived success and intercausal perceptions. Finally, Porac suggested that the relationship between perceived intercausal influence and causal attributions suggest that intercausal perceptions are involved in the more general attribution process.

Porac's study demonstrated that causal attributes are often linked in a reciprocal fashion into a causal loop. In a causal loop, one variable has inputs to a second variable and this latter cause loops back to influence the first. A reciprocal interaction is thus set in motion. For example, someone who has not expended much effort in studying for a test might perceive the test as being quite difficult. In this instance, lack of effort has directly influenced perceived task difficulty, which in turn influences judgments of effort. In such instances, a causal loop has formed.

Besides identifying causal attributes of past performance outcomes and their intercausal relationships, attribution theory has also attempted to identify an underlying structure of perceived causality. The reason behind these attempts is to develop a taxonomy

for classifying and comparing various causal attributes (Weiner, 1985).

One dimension of attribution theory is an internal-external (locus) dimension. Since the 1950's, according to Collins, Martin, Ashmore, and Ross (1974), psychologists have acknowledged an internal-external distinction. Internal factors are those found within a person and external factors represent those factors found within the environment (Heider, 1958).

Stability, a second dimension of causality, was identified by Weiner, Frieze, Kukla, Reed, Rest and Rosenbaum (1971). The reasoning behind adding stability as another dimension was that some internal and external causes fluctuate, whereas others remain fairly constant. By integrating this second dimension, causal attributes may be viewed as being either internal or external, and as either stable or unstable. For example, ability can be classified as internal and stable, effort as internal and unstable, task difficulty as external and stable, and luck as external and unstable.

Weiner (1983) later questioned the validity of this second dimension of stability. He argued that ability may be perceived as unstable if learning is possible. Effort may be viewed as a stable trait, as when we describe the personalities of an individual using the labels lazy or industrious. Weiner further noted that tasks can be changed to be more or less difficult and luck may be thought of as a property of an individual. Someone is either lucky or unlucky.

A third dimension, controllability, was also identified (Weiner, 1979). Rosenbaum (1972) originally recognized that some causal

attributes, such as mood, fatigue, and temporary effort, could be subject to volitional control. That is, an individual can increase or decrease the effort expenditure concerning these attributes. Controllability then, was thought to add yet another dimension to causal attributions. Weiner (1985) suggests that the concept of controllability is an important contributor to emotions. When we can control our behavior and don't, we frequently feel guilty, whereas if we are unable to control unwanted behaviors, we feel shame. Conversely, if we view someone as being able to control unacceptable behaviors and they don't, we often express anger. Pity, however, is frequently our emotional reaction to unacceptable behavior that cannot be controlled.

To summarize this section, attribution theory attempts to identify the causal factors thought by individuals to explain performance outcomes. Numerous investigations reveal that people commonly select the same few attributes in explaining performance outcomes. Effort, ability, task difficulty, and luck have been repeatedly identified as major causal factors in the achievement domain. Causal factors of performance outcomes may be classified along three dimensions; locus (internal- external), stability (stable-unstable), and control (controllable- uncontrollable). Finally, an individual's attributions concerning past performance or failure may exert important effects on self-efficacy judgments. Such attributes (ie., ability, effort, task difficulty, and luck) may represent potential sources of efficacy information.

As suggested, individuals may perceive past performance

outcomes in terms of causal factors. If for example, they exerted a great deal of effort in studying for a college examination and received a high grade on their exam, they may link the amount of effort exerted to exam outcome. Any self-prediction for subsequent test performance would take into account the amount of effort exerted in studying. If little effort was expended, a weaker efficacy judgment would most likely be made.

Problems with this line of reasoning may occur, however, if the individual has incorrectly linked performance outcome with a specific causal factor. Perhaps in the above cited example, effort was not the predominant causal factor. If this were indeed the case, subsequent test prediction estimates would most likely be misjudged. It is suggested by the author that a correct understanding of the causal factors affecting one's own past performances is an important criterion in the accurate prediction of future performance outcomes.

Section Four: Gender and Sexual-Role Identity

Studies exploring gender effects on self-efficacy judgments have been equivicol. Some studies have identified significant gender interactions, while many others have not. For example, Campbell and Hackett (1986), in a study on the effects of mathematics task performance on math self-efficacy for college undergraduates, found significant gender differences. Specifically, they found women rating themselves significantly lower on strength of self-efficacy measurements than did men.

Taylor (1985) found that although self-efficacy was a

significant predictor of sports performance for both male and female varsity athletes, self-efficacy accounted for greater variance in sports performance for females than it did for males. Taylor could find no reasonable explanation to interpret these findings and suggested further investigation into the affects of gender on self-efficacy judgments was needed.

Numerous other studies (Gist, 1987), however, have found no significant gender interaction. Studies by Lent, Brown, and Larkin (1986) and Lent and Larkin (1984) investigating the role that self-efficacy plays in academic performance of college undergraduates found no significant gender interactions.

The role that gender may exert in determining the accuracy of self-predictions is still not well understood. Conflicting results indicate that gender may have an effect on self-predicted performance outcomes, although its exact role is unknown. A hypothesis of this study is that it may not be gender per se that is mediating self-prediction accuracy but rather perceived sex-role identity, which is non-gender specific.

Bem (1981) has taken a different approach in her research on gender, concentrating on the control that sex-role identity exerts on human behavior irrespective of actual gender affiliation. Although her work has not included self-efficacy, never-the- less, her concepts may help explain some of the variance attributed to gender in self-efficacy research.

Traditionally, sex-role identity has been viewed as either masculine or feminine, depending upon one's gender. Bem, however,

views sex-role identity as representing behavioral traits irrespective of gender. Masculine behavioral traits according to Parsons and Bales (1955) are associated with an instrumental orientation and a cognitive focus on getting the job done. Feminine traits are associated with an expressive orientation and an affective concern for the welfare of others and the harmony of the group.

Bakan (1966) has made similar observations, suggesting that masculinity is associated with an agentic orientation and a concern for oneself as an individual. Femininity suggests Bakan, is concerned with a communal orientation, representing a concern for the orientation of oneself and others.

High masculinity during adulthood has been correlated with high anxiety, high neuroticism, and low self-acceptance. (Hartford, Willis, and Deabler, 1967). Conversely, high femininity has consistently been correlated with high anxiety, low self-esteem, and low social acceptance (Cosentino and Heibrun, 1964; Gall, 1969; Gray, 1957; Sears, 1970; Webb, 1963). Macoby (1966) has found that boys and girls who are more strongly sex-typed have been found to have lower overall intelligence, lower spatial ability, and lower creativity.

Bem's (1981) research has focused on the concept of psychological androgyny. Androgyny assumes, at least in principle, that an individual may exhibit both masculine and feminine behavioral traits. Such an individual, according to Bem, may be both expressive and instrumental, depending upon the situational appropriateness of these two modalities.

The concept of psychological androgyny is not unique to Bem.

Jung's (1953) theory described the presence of the anima and the animus, which was thought to be present in all humans. Bakan (1966) has also argued for the positive benefits, both to the individual and society, of possessing both agency and communion.

In relation to self-efficacy judgments, is it possible that sexual-role identity, and not gender, may significantly affect the accuracy of self-predictions? Are self-efficacy judgments negatively affected by individuals high in either masculinity or femininity traits? Would androgenous individuals be better predictors of personal ability than either of the two extremes? These questions appear relevant to any study of self-efficacy judgments, and may assist in explaining some of the significant gender interactions in previous self-efficacy research.

Individuals exhibiting predominantly masculine identities may tend to overestimate their own abilities. Conversely, individuals of only high feminine traits may tend to undersetimate their own abilities. In both cases, unrealistic and inaccurate self-predictions would result. An androgenous individual, however, may balance these two extremes.

Section Five: The Type A - Type B Personality Behavior

As noted by Ward and Eisler (1987), the accuracy of self-predictions is affected by the Type A - Type B behavior pattern continuum. The Type A behavior personality refers to a competitive, multiphasic, achievement oriented person who is impatient, easily aroused, hostile, and angry (Rosenman, Friedman, Strauss, Wurm,

Kositchek, Hahn, and Werthessen, 1964). According to Wright (1988), three traits especially characterize the Type A behavior personality; a sense of time urgency, a multiphasic orientation, and chronic activation. A sense of time urgency refers to a preoccupation with saving small amounts of time, usually measured in seconds. A multiphasic orientation identifies the need to undertake multiple projects or do more than one thing at a time. Chronic activation refers to staying active or keyed up for most of the day.

A Type B personality behavior was originally defined (Rosenman, Friedman, Strauss, Wurm, Kositchek, Hahn, and Werthessen, 1964) as someone not exhibiting Type A personality traits. Friedman and Rosenman (1974) later argued that Type B individuals may be just as ambitious and achievement oriented as their Type A counterparts. The ambition associated with Type B individuals, however, is characterized by confidence and satisfaction, whereas the ambition associated with the Type A behavior pattern is dominated by anxiety and anger. A third personality type, the Type AB, is used by some authors (Ward and Eisler, 1987) to denote an individual exhibiting both Type A and Type B personality characteristics. These three personality behaviors are commonly assessed using structured interviews or self report questionnaires (Mathews, Krantz, Dembroski, and MacDougall, 1982).

Much of the research concerning the Type A - Type B personality behavior pattern originated within the medical community. The reason behind this interest is well summarized in the findings of the National Institute of Health's Review Panel on

Coronary Prone Behavior and Coronary Heart Disease (1981) when they stated, " the available body of scientific evidence demonstrates that Type A behavior is associated with an increased risk of clinically apparent CHD (chronic heart disease) in employed middle-aged U.S. Citizens. This increased risk is greater than that imposed by age, elevated levels of systolic blood pressure, serum cholesterol, and smoking" (p. 1200).

According to Wright (1988), three factors play a critical role in the development of the Type A personality behavior prone to chronic heart disease. First is a high need to achieve. Second is a history of early success and subsequent reinforcement for striving efforts. The third factor is an exposure to timed activities that provide a personal blueprint for achieving more by efficiently managing time and by chronic activation. Wright labels these three factors as predisposing.

Although they may be potentially dangerous, the three factors become lethal according to Wright only when the Type A individual also exhibits low self-esteem. Apparently in an effort to raise self-esteem, the Type A individual attempts to achieve more and more. Unfortunately this approach commonly invites failure, which appears to only heighten the need to set even more goals of greater difficulty.

Friedman and Rosenman (1974) capture this vicious cycle when they state that the Type A personality behavior is, "above all a continuous struggle, an unremitting attempt to accomplish or achieve more and more things" (p. 31). Research appears to support Friedman and Rosenman's position, suggesting that in achievement situations,

Type A individuals, when compared to their Type B counterparts, tend to be more competitive (Van Egeren, 1979) and more hard driving (Weidner and Mathews, 1978). Type A individuals also tend to set difficult performance goals for themselves. Grimm and Yarnold (1984) and Price (1982) suggested that Type A individuals set excessively high and inflexible standards for their own performance.

Ward and Eisler (1987), in two separate experiments, found that Type A individuals are less likely to achieve personal goals than are Type B or Type AB individuals. Throughout the two experiments conducted, Ward and Eisler found Type A individuals repeatedly setting goals in excess of their actual performance potential. In contrast, Type B and Type AB individuals consistently underestimated or correctly predicted their performance potential. Unfortunately, Ward and Eisler did not specifically identify those individuals who underestimated and those individuals who accurately predicted their performance potential. The two authors concluded their studies by suggesting that the Type A behavior pattern is associated with a low probability of achieving predicted performance.

The reviewed research on the Type A - Type B personality behavior continuum indicates that such predispositions can significantly affect the accuracy of efficacy judgments. The literature suggests that Type A individuals consistently overestimate their performance capabilities. It is somewhat less clear, however, on the effects of the Type AB and Type B personality behaviors. How the Type A - Type B behavior continuum influences the accuracy of efficacy judgments is not known. One suggestion is that behavior type

influences the types of efficacy information cues selected and/or weighted. This hypothesis, however, has not been empirically tested.

Section Six: Summary

Chapter II provided a review of the literature pertaining to the current study. The concept of self-efficacy was initially defined and the four important sources of efficacy information (enactive experiences, vicarious experiences, verbal persuasion, and physiological state) proposed by Bandura (1982) discussed. Guided by Schunk's (1985) suggestion that causal factors of human performance developed from attribution theory may also represent important efficacy cues, the theory of attribution was summarized and commonly identified attributes in the achievement domain cited.

The effects of gender on self-efficacy judgments were also reviewed. The equivicol findings suggested that gender may not be a significant factor in determining the accuracy of efficacy judgments. Rather the work of Bem (1981) on sex-role identity, irrespective of gender, may prove to be a more promising research avenue.

Finally, the effects of the Type A - Type B personality behavior continuum on the accuracy of efficacy judgments was discussed. Research has documented that Type A individuals frequently overestimate their own performance potential. Conversely, their Type B and Type AB counterparts seem to either underestimate or accurately predict their performance potential (Ward and Eisler, 1987). The psychological reasons for such differential predictive accuracy are unclear, however.

The reviewed literature certainly substantiates the premise that judgments of self-efficacy can be reliable predictors of performance outcomes in a variety of domains if the efficacy measure is adapted to the specific tasks being measured. The literature is less clear when one attempts to understand how efficacy judgments are made. No studies have focused specifically on identifying the types of information sources people select in making self-predictions or how such information is ultimately weighted and integrated into a final efficacy judgment. Further, little is known about why some people can consistently make accurate predictions of their own performance while others can't.

One avenue of investigation, and the one chosen for this study, is to assess the possibility that people select and weight efficacy cues differently (Bandura, 1986 and Schunk, 1985). Such differential treatment of efficacy information may account for the variance in the accuracy of self-prediction. This remark carries with it the assumption that some cues may be more critical to the accurate prediction of personal performance than other cues. Bandura (1986) alludes to this concept when he suggests that people who focus on their physiological state when developing efficacy judgments may consistently underestimate their abilities.

Using an attributional theory approach, it would seem that individuals who explain past performance outcomes in terms of luck or other highly unstable variables would be at a disadvantage in making accurate self-predictions. Conversely, individuals attributing task performance to effort and ability may hold an advantage when

making predictions of their own performance. These ideas, however, have not been tested.

One of the basic differences between attribution and self-efficacy theory is the types of information cues selected. Bandura (1982) pays particular attention to past performance, either as directly experienced by the individual or vicariously. In contrast, attribution theory places more importance on specific task preparation and execution, examplified by the attribute of task effort. Kahneman and Tversky (1973), in their work on the psychology of intuitive predictions, suggests that more accurate predictions are made when individuals consider past performance or what they term the statistical base rate. Their studies demonstrated, however, that individuals rarely use this information base, even when readily available. Instead, study participants almost always selected task specific information. Kahneman and Tversky termed this judgmental bias representativeness.

It is suggested that the accuracy of efficacy judgments is dependent upon a number of variables. The exact interrelationship of these variables, however, is not understood. It is unclear whether specific personality traits such as anxiety level and Type A - Type B behavior pattern, directly influence the types of efficacy information cues selected and/or weighted. Although such a hypothesis may be intuitively attractive, it has not been empirically tested.

CHAPTER III

METHODOLOGY

Chapter III describes the specific methodologies used in the study and is divided into five sections. Topics covered include characteristics of the subjects participating in the study (Section one), the instrumentation used in gathering the data (Section two), the specific procedures followed in collecting the data (Section three), and the the research design (Section four).

Section One: Subject Characteristics

Subjects participating in the study were 157 university undergraduate students majoring in education. All students were enrolled in the same required undergraduate educational psychology course at a large, mid-western, public university. Three separate sections of the course, each taught by a different instructor, participated in the study. Each section used the same syllabus, the same required text, the same assignments, and the exact same four, 40-point, multiple-choice tests.

Of the 157 participants, 111 (71%) were female and 46 (29%) were male. Subjects ranged in age from 19 to 43, with a mean age of 23.4 and a standard deviation of 5.1. Three percent of the students were sophomores, 42% were juniors, 51% were seniors, and 4% were graduate students. Cumulative undergraduate GPA of the students ranged from 2.0 to 3.92, with a mean GPA of 3.04 and a standard

deviation of 0.45.

Participation in the study was voluntary. Students wishing to participate received extra credit. Those students who selected not to participate in the study were also given an opportunity to receive equal extra credit by another means.

Section Two: Instrumentation

Four instruments were used in the study. The first was a questionnaire developed specifically for this study to investigate information sources selected and weighted by individuals in making efficacy judgments. The instrument, termed the Exam Prediction Questionnaire, will be referred to by using the acronym EPQ. The second instrument employed was the Bem Sex-Role Inventory (Bem, 1981). The Bem Sex-Role Inventory is commonly used to determine perceived sex-role identity. It will be referred to by using the acronym BSRI. The third instrument used was the Test Anxiety Inventory (TAI) developed by Spielberger (1980). The TAI measures trait test anxiety of students. The final instrument employed was the Jenkins Activity Survey (JAS), Form T (Kantz, Glass, and Snyder, 1974). The JAS measures the Type A - Type B personality behavior continuum.

Exam Prediction Questionnaire

The exploratory nature of this study and the fact that little research has been conducted concerning how individuals select and weight efficacy information cues for exam performance precludes the use of any existing instrument. As a result, an original instrument,

referred to as the EPQ, had to be developed. The following discussion of the EPQ will be divided into three parts. The first part reviews the work of Elig and Frieze (1979). Some of Elig and Frieze's work was summarized in Chapter II, however, their validity studies of instrumentation used in attribution research appears especially pertinent to this discussion of the EPQ and consequently are summarized here rather than in Chapter II. The second part reviews the development and content of the EPQ. The final part documents pilot studies involving the EPQ and list subsequent changes to the instrument as a result of this initial testing.

Work of Elig and Frieze (1979). As noted in Chapter II, attribution theory attempts to identify the causal factors thought by individuals to explain performance outcomes. It differs from self-efficacy judgments in that attributions are post- performance whereas self-efficacy judgments are pre-performance. As suggested by Schunk (1985), however, an individual's attributions concerning past performance or failure may exert important effects on subsequent self-efficacy judgments. Weiner (1979) has also noted that future performance expectancies such as self-efficacy may depend heavily on causal ascriptions. As a result of the similarity between self-efficacy and attribution theory, close attention was given to the design and use of instrumentation in the area of attribution theory.

A number of research articles concerning causal attributions for success and failure were published in the 1970's and 1980's (see Chapter II for a comprehensive summary). As noted by Elig and Frieze (1979), however, little of this research investigated how causal

attributions should be measured. In an attempt to rectify this situation, the two authors investigated the interrelationship of several measures of causal attributions to assess their validity and to offer recommendations concerning the selection of instruments to be used in future research.

Elig and Frieze (1979) described a number of different techniques commonly employed in assessing causal attributes. These techniques may be grouped into two major categories. One category involves open-ended responses and the second one involves more structured responses, such as independent ratings, ipsative ratings, choice of one major cause, and bipolar ratings. Open-ended responses ask subjects to state in their own words why a particular event has occurred. The verbal responses are then classified by a skilled rater into any set of previously defined attributional categories. A positive aspect of using open-ended responses is that they allow subjects to mention attributes that may not have been identified earlier by the researcher. Two major limitations of using open-ended responses involve the necessity of training coders and the time-consuming nature of this type of causal assessment.

According to Elig and Frieze, a major distinction between various structured attribution measures is whether the responses involve ipsative or independent judgments. Ipsative measures are measures in which the score of one attribution, by definition, must influence the score of the other attributions, thus inducing negative correlations (Elig and Frieze, 1979). Among the ipsative measures, the assignment of percentages to various attributional causes is

perhaps the most widely used. Elig and Frieze noted that using percentage ratings makes explicit the basic assumption of independent judgments because, "the causes being rated account for the totality of cause for the outcome and that the total cause of an event can be parceled out to various particular causes" (p. 623).

Negative correlations are not forced by measures involving independent ratings. The use of independent ratings in attribution research has been quite prevalent (Elig and Frieze, 1979). When using independent ratings, subjects are normally asked to rate some particular attribute using a Likert-type scale. Scales commonly range from a low of 1 to some higher number, usually 5, 7, or 9. One of the major strengths of independent ratings is that they offer ease of analysis because each attribution may be tested separately. More specifically, this procedure yields quantitative data rather than nominal.

A major problem with the use of structured responses according to Elig and Frieze (1979) is that it confines subjects to a limited set of factors. These factors have been defined in advance by the experimenter as being important for the particular situation under study. This a priori set, however, may not include the factors of importance for some subjects. The repeated use of only four attributional causes (effort, ability, luck, and task difficulty) in many structured responses involving attributional research seems to have exacerbated this problem.

In their research, Elig and Frieze (1979) examined the validity issue concerning various attributional measures. Specifically, they

compared open-ended response measures with two structured response measures, one using independent judgments and the other using ipsative judgments. The independent judgment measurement used a 7-point Likert scale.

Hypotheses for their study included: (1) convergent and discriminant validities will be lower for the open-ended response measure than for either of the two structured response measures, (2) the face validity to subjects of the open-ended response question will be better than that obtained by either structured response measures, and (3) the independent ratings of the structured response will be superior to the percentage ratings in terms of convergent and discriminant validities.

Participants in the study were college undergraduates. After completing a series of anagrams, students expressed attributions concerning their performance outcomes using the three different attribution measures. After completing all three forms, subjects were asked a series of questions concerning the three instruments they had just used. Results from the study supported the first hypothesis.

Structured response reliabilities were higher than those of open-ended responses. Convergent and discriminant validities for structured measures were found to be satisfactory, whereas open response convergent validities were quite low.

The second hypothesis, that the face validity of the open-ended responses would be higher than the two structured responses, was not supported. Instead college students preferred open-ended responses

and independent ratings and disliked the ipsative rating scale which employed a percentage method. According to Elig and Frieze (1979), "subjects said they felt that the percentage measure was hard to compute and was not the best reflection of what they felt were the reasons for the outcome" (p. 631).

The third hypothesis was supported. Independent ratings of the structured response were significantly superior to the percentage ratings in terms of convergent and discriminant validities.

Elig and Frieze (1979) concluded that, at least for college students, independent ratings using a scale method is clearly the "superior technique." Elig and Frieze further suggested that independent ratings could be improved if future researchers provide subjects with a wider selection of causal factors from which to choose.

As a result of Elig and Frieze's findings concerning university student preference for an independent rating scale, a 7-point Likert scale was chosen for the EPQ. The Likert scale was used to estimate the perceived influence each efficacy information source exerted in making a predicted test score. Rankings ranged from 1 (of no influence) to 7 (of extremely high influence). Also, Elig and Freeze's suggestion that a wide range of factors should be included on any instrument was adopted. The EPQ contained seventeen efficacy information sources, allowing respondents a wide range of options from which to choose.

Content of the EPQ. As shown in Appendix A, the EPQ is a single-sheet, double-sided questionnaire containing five separate questions. Question #1 asks respondents to predict how many

multiple-choice questions they will answer correctly on the 40-point examination that they are about to take. Question #1 measures the level of an individual's efficacy judgment. Subjects may select only one number between 0 and 40.

Question #2 asks subjects to state how confident they are in the prediction they made in Question #1. A confidence range from 0% to 100% in ten point increments is provided. Question #2 is a measure of the confidence of prediction.

Question #3 asks subjects to rate independently the degree of influence that each of 17 information sources exerted while they were making their test score prediction. A 7-point Likert scale is used ranging from 1 ("of no influence") to 7 ("of extremely high influence"). This method for assessing information cue selection and weighting was chosen based primarily on the research of Elig and Frieze (1979) which was previously discussed.

To improve the content validity of the EPQ, all information cues were selected from past studies. These studies were described in Section three of Chapter II. Information cues identified from studies involving attribution research as a perceived cause of academic achievement or from self-efficacy research involving the identification of sources of information used in making efficacy judgments were included in the EPQ. The seventeen factors included in the EPQ and references to support their inclusion are:

1. Teacher effectiveness (Bar-Tal, Goldberg, and Knani, 1984).

2. Other students performance in the course (Bandura, 1977;

Bandura, Adams, Hardy, and Howells, 1980; Kazdin, 1974).

3. Perceived test difficulty (Elig and Frieze, 1979; Frieze and Snyder, 1980; Weiner, 1976 &1979; Wilson and Palmer, 1983).

4. Past performance in other courses (Bandura, 1977; Bandura, Adams, Hardy, and Howells, 1980).

5. Self-confidence level (Bar-Tal, Goldberg, and Knani 1984).

6. Concentration level while studying (Burger, Cooper, and Good 1982; Cooper and Burger, 1980).

7. General academic ability as a student (Anderson, 1983; Bar-Tal, Goldberg, and Knani, 1984; Burger, Cooper, and Good, 1982; Cooper and Burger, 1980; Frieze, 1976; Elig and Frieze, 1979; Frieze and Snyder, 1980; Parkerson, Lomax, Schiller, & Walberg (1984); Walberg (1981); Weiner,1976 &1979; Wilson and Palmer, 1983).

8. Anxiety level (Bandura and Adams, 1977; Hunsley, 1985; Naveh-Benjamin and McKeachie, 1987; Paulman and Kennelly, 1984).

9. Knowledge of material to be tested (Anderson, 1983; Bar-Tal, Goldberg, and Knani, 1984; Burger, Cooper, and Good, 1982; Cooper and Burger, 1980; Frieze, 1976; Elig and Frieze, 1979; Frieze and Snyder, 1980; Parkerson, Lomax, Schiller, and Walberg (1984); Walberg (1981); Weiner,1976 and 1979; Wilson and Palmer, 1983).

10. Past performance in this course (Bandura, 1977; Bandura, Adams, Hardy, and Howells, 1980).

11. Interest in test material (Frieze and Snyder, 1980; Wilson and Palmer, 1983).

12. Present mood (Elig and Frieze, 1979).

13. Encouragement given by teacher (Bandura, 1982).

14. Physical health (Schunk, 1985).

15. Study skills (Anderson, 1983; Bar-Tal, Goldberg, and Knani, 1984; Naveh-Benjamin and McKeachie, 1987).

16. Luck (Frieze, 1976; Weiner, 1976 & 1979; Wison and Palmer 1983).

17. Effort devoted to studying (Anderson, 1983; Bar-Tal, Goldberg, and Knani, 1984; Burger, Cooper, and Good, 1982; Cooper and Burger, 1980; Frieze, 1976; Elig and Frieze, 1979; Frieze and Snyder, 1980; Parkerson, Lomax, Schiller, and Walberg (1984); Walberg (1981); Weiner,1976 &1979; Wilson and Palmer, 1983).

In Question #4, the seventeen information cues are listed again. This time, however, subjects are asked to select only the four most important sources of information used in making their test

predictions. Once these four sources have been selected, subjects are further asked to rank-order them from 1 (most important) to 4 (least important).

Weiner (1985) has repeatedly maintained that there are four major causal attributions for success/failure: effort, ability, task difficulty, and luck. Bandura (1982) suggested that there are four major sources of efficacy information: enactive experiences, vicarious experiences, verbal persuasion, and physiological state. Question #4 was constructed to determine if Weiner's four causes would be chosen more frequently than Bandura's four sources, some combination of Weiner's and Bandura's categories chosen, or other information cues selected. Also, question #4 would be used to identify differences in information sources selected by individuals who accurately predict their test scores from those who do not.

Question #5 elicits demographic information about the subject. Age, gender, class, cumulative grade point average, educational major and name were requested.

Pilot Testing. Three pilot tests of the EPQ were conducted. The EPQ was initially pilot tested using 33 college students majoring in education and enrolled in a required education course similar to the one that was finally used during the study. Students were first given the verbal descriptions of each scale point in question #3 in a random order and asked to rank order them from 1 (lowest) to 7 (highest). The initial seven descriptions used on the EPQ for the pilot study were: (1) of no influence, (2) of very slight influence, (3) of slight influence, (4) of moderate influence, (5) of high influence, (6) of very

high influence, and (7) of critical influence.

Results from this portion of the pilot study indicated that 15% of the students did not fully understand the meaning of the word "moderate". Also, 32% of the students did not correctly rank #7 ("Of critical influence"). As a result, "moderate" was changed to "medium" and "of critical influence" was changed to "of extremely high influence".

After students independently ranked the descriptors, the EPQ was distributed. Students were asked to pretend that they were about to take a 40-point, multiple-choice exam and asked to fill out the questionnaire accordingly. Students were also asked to underline all words, phrases, or sentences in the questionnaire that were not perfectly clear to them.

After the students had completed the questionnaire, each section of the EPQ was read aloud. Students were asked if any phrasing was unclear. Students were also randomly selected and asked to interpret in their own words what a particular word or phrase used in the questionnaire meant. After the student voiced his or her opinion, other students were asked if they concurred. Following this portion of the pilot study, all questionnaires were collected. Each questionnaire was carefully examined to determine if the student correctly filled it out and to note any underlined words or phrases. As a result of this initial pilot testing, a number of changes to the questionnaire were made. Perhaps the biggest problem identified during the first pilot test was the confusing wording used in the instructions for Question #4. Four students incorrectly filled

out Question #4. When asked the reason, all four students replied that the instructions accompanying Question #4 were not clear.

After the suggestions of the students were incorporated into the EPQ, two further pilot tests were conducted. The same procedure was followed that was employed during the first pilot study. In the second pilot test, 43 college students majoring in education were used from another required educational course. No major problems were identified with the questionnaire, although some minor wording was changed. The split-half reliability estimate for this second pilot test, based only on parts 3 and 4, was .91.

The third pilot study involved 36 college students majoring in education from another section of the course used in Pilot Study #2. Once again, the same procedure was used as in the two previous pilot studies. During this final pilot test, no problems were identified. The split-half reliability estimate for this third pilot test, based only on parts 3 and 4, was .87.

The final edition of the EPQ is shown in Appendix A. The questionnaire takes approximately 7 to 10 minutes to complete.

The Bem Sex-Role Inventory

Femininity and masculinity have been normally conceptualized as opposite ends of a single bipolar dimension. Recently, however, researchers in a number of disciplines have focused on the concept of psychological androgyny (Bem, 1981). According to Bem (1981), psychological androgyny denotes, "the integration of femininity and masculinity within a single individual" (p. 4). Psychological androgyny

implies that it is possible, at least theoretically, for an individual to be both compassionate and assertive, both expressive and instrumental, both feminine and masculine, depending upon the situational appropriateness of these various modalities (Bem, 1981).

Content of the BSRI. The Bem Sex-Role Inventory (BSRI) was developed to implement empirical research on psychological androgyny. The BSRI contains 60 personality characteristics. Twenty of the characteristics are stereotypically feminine (e.g., affectionate, gentle, understanding, sensitive to the needs of others) and twenty are stereotypically masculine (e.g., ambitious, self-reliant, independent, assertive). Twenty items that serve as filter items (e.g., truthful, happy, conceited) are also included in the BSRI. Items selected for the BSRI were initially judged by 100 (50 males and 50 females) undergraduate students at Stanford University in 1972 (Bem, 1981).

When taking the BSRI, subjects are asked to indicate on a 7-point scale how well each of the 60 characteristics describes themselves. The scale ranges from 1 ("Never or almost never true") to 7 ("Always or almost always true") and is labeled at each point. The BSRI takes approximately 10 - 15 minutes to complete.

According to Bem (1981), the BSRI has two features that distinguish it from most masculinity-femininity scales. The first feature, and perhaps most important, is that the BSRI treats femininity and masculinity as two independent dimensions rather than as two ends of a single dimension. This feature allows an individual to indicate whether he or she is high on both dimensions

("androgynous"), low on both dimensions ("undifferentiated") or high on one dimension but low on the other (either "feminine" or "masculine").

The second feature of the BSRI involves the nature of the items. All items were selected as feminine or masculine on the basis of cultural definitions of sex-typed social desirability and not on the basis of differential endorsement by males and females.

Psychometric Analyses. Psychometric data were collected for the BSRI from two samples of subjects, both consisting of undergraduate students in introductory psychology courses at Stanford University. The first sample included 279 females and 444 males who filled out the BSRI in 1973. The second sample included 340 females and 476 males who completed the BSRI in 1978.

In order to estimate the internal consistency of the BSRI, coefficient alpha was computed separately for females and males in both samples for the Femininity score, the Masculinity score, and the Femininity minus Masculinity Difference score. Derived coefficient alphas varied from a low of .75 to a high of .87. In order to examine test-retest reliability, the BSRI was adminis- tered for a second time to 28 females and 28 males from the 1973 Stanford sample. The second administration took place approximately four weeks after the first. During the second administration, subjects were explicitly told not to try and remember how they had responded previously. Product-moment correlations were computed between the first and second administrations. Reliability scores ranged from a low of .76 to

a high .94.

To check the relationship between social desirability response-set and an individual's scores on the BSRI, the Marlowe-Crowne Social Desirability scale was administered along with the BSRI to the 28 females and 28 males in the 1973 test-retest sample. Product-moment correlations were computed between the two instruments. All the correlations were quite low, ranging from -.15 to .21. Bem (1981) concluded from these data that the BSRI scores are not measuring a general tendency to describe oneself in a socially desirable manner.

Test Anxiety Inventory

As discussed in Chapter II, Bandura (1982) hypothesized that anxiety level may represent a poor source of efficacy information. According to Bandura, individuals who focus on their own anxiety prior to a task may significantly underestimate their abilities and expected performance attainments. This hypothesis, however, has not been empirically tested. To test Bandura's hypothesis, all subjects in the study were given the Test Anxiety Inventory (Spielberger, 1980) or TAI, a self-report psychometric scale used to measure individual differences in anxiety proneness to test situations.

Content of the TAI. The TAI was developed to measure individual differences in test anxiety as a situation-specific personality trait (Spielberger, 1972). The test form is one page and includes directions, twenty items, and space for recording responses. Respondents are asked to report how frequently they experience

specific symptoms of anxiety before, during, and after examinations. The inventory is similar in concept and structure to the A-Trait Scale of the State-Trait Anxiety Inventory (STAI), which measures general anxiety proneness in adolescents and adults (Spielberger, Gorusch, and Lushene, 1970).

The construction and development of the TAI was guided by the concepts of worry and emotionality. Liebert and Morris (1967) identified worry and emotionality as the two major components of test anxiety. They defined worry as cognitive concerns about the consequences of failure and emotionality as reactions of the autonomic nervous system that are evoked by evaluative stress. According to Spielberger, Gonzalez, Taylor, Algaze, and Anton (1978), worry and emotionality may also be thought of as major components of the state-trait anxiety reactions experienced in test situations, including tension, apprehension, nervousness, and arousal of the autonomic nervous system.

The TAI was originally developed to measure test anxiety in high school and college students. As described by Spielberger (1980), the TAI was designed for self-administration and may be given individually or in groups. Although there are no time limits, most high school and college students complete the inventory in eight to ten minutes.

While taking the inventory, respondents use a four-point scale to report how frequently they experience specific symptoms of anxiety in test situations. The four choices are: (1) almost never, (2) sometimes, (3) often, and (4) almost always. For example, in response

to item 15, "I feel very paniky when I take an important test," the students select the response that best describes how they generally feel during tests.

All twenty items are used to determine the TAI total score. Since each response may be weighted from one to four, the minimum TAI total score is 20 and the maximum is 80.

Psychometric Analyses. Normative data of the TAI are based on studies of large samples of college undergraduates, entering college freshmen, and high school students (Spielberger, 1980). The TAI norms for college students are based on 1,449 undergraduates (654 males and 795 females) and 1,129 incoming freshmen (533 males and 596 females) from the University of South Florida. The undergraduates graduates were given the TAI in introductory psychology courses and the freshmen were tested during a summer orientation program immediately prior to their first year of college.

Test-retest reliabilities were determined for time periods of 2 weeks, 3 weeks, 4 weeks, and 6 months. For the shorter periods (2 - 4 weeks), the reliability coefficients were .80 or higher, but dropped to .62 for the 6 month period (Spielberger, 1980). Validity for the instrument was established by correlating the TAI with six other anxiety measures. Of special note is the high correlation with Sarason's (1978) Test Anxiety Scale (TAS). The correlation of the TAI total score with the TAS was .82 for males and .83 for females. Spielberger (1980) concluded from his analysis that the 20-item TAI total score and the 37-item TAS are essentially equivalent measures.

Jenkins Activity Survey, Form T

Studies by Grimm and Yarnold (1984), Price (1982), and Ward and Eisler (!987) demonstrated that efficacy judgments may be affected by the Type A - Type B behavior continuum. All three studies found that Type A individuals commonly set excessively high and inflexible standards for their own performance. Ward and Eisler concluded their studies by suggesting that the Type A behavior pattern is associated with a low probability of achieving predicted performance.

In an attempt to study the relationship between the Type A - Type B behavior pattern continuum and cumulative accuracy score, the Jenkins Activity Survey (JAS), Form T (Krantz, Glass, and Snyder, 1974), was administered to all study participants. As reported by Glass (1977), the JAS is ideally suited to be given to large groups of people when lengthy personal interviews are simply not possible.

Content of the JAS. Initially the Type A - Type B behavior pattern was assessed by a standard behavioral interview known as the Structured Interview. In the structured interview, behavior pattern classification was based upon subjective clinical judgments by trained raters. In a large scale study known as the Western Collaborative Group Study (WCGS), individuals classified as Type A by the interview method were observed to have roughly twice the incidence of chronic heart disease compared to their Type B counterparts (Blumenthal, 1985).

According to Blumenthal (1985), the Jenkins Activity Survey (Jenkins, Zyzanski, and Rosenman, 1971) was developed in an attempt to duplicate the clinical assessment of the Type A - Type B behavior

pattern by employing an objective psychometric procedure. The JAS is a self-administered, multiple-choice questionnaire that yields a composite Type A - Type B score based on the scoring of twenty-one of the forty-four questions. A typical question on the JAS would be "Has your spouse or some friend ever told you that you eat too fast?" A Type A response to this question is "Yes, often" whereas a Type B response would be "Yes, once or twice" or "No, no one has told me this."

The JAS was originally developed for employed middle-class males. Administration of the JAS to a college student population, therefore, is not entirely appropriate. Recognizing this fact, Krantz, Glass, and Schaeffer (1974) modified the wording of several items in the JAS to reflect this orientation. This modified version has become known as the JAS, Form T, and has been widely used in studies involving university students (Ward and Eisler, 1987).

Glass (1977) described the modification of the JAS for student use. According to Glass, items in the original JAS referring to income, job involvement, and job responsibility were either eliminated from or modified for the student version of the questionnaire. The word "courses" for example, was substituted for the word "job."

The student version of the JAS is scored by a unit-weighting procedure (Glass, 1977). For each of the 21 items on the A - B scale that are scored, the A responses receive a 1 and the B responses receive a score of 0. According to Glass, the median A-B score for college males typically falls between 7 and 8, and for college females between 6 and 7, where 0 is the maximal Pattern B score and 21 is

the maximal Pattern A score. A scoring system of 0 to 5 for Type B individuals and greater than 10 for Type A individuals was adopted by Krantz, Glass, and Schaeffer (1974).

Psychometric Analysis. The original normative data for the JAS was based on 2,588 employed middle class males age 48 through 65 who participated in the Western Collaborative Group Study (Blumenthal, 1985). Derived coefficient alphas varied from a low of .42 to a high of .85. Much of the validity efforts attempted to correlate the JAS with the structured interview. Test - retest reliabilities for the original sample ranged between .60 and .70 over periods from six months to four years.

Glass (1977) established test - retest reliabilities for the student version. In one experiment involving 459 university students from Texas, test - retest reliabilities were .85 and higher for time periods ranging from 2 weeks to 4 months. Also using the student version, Nielson and Dobson (1980) demonstrated strong support for the discriminant validity of the Type A behavior pattern in relation to trait anxiety.

Section Three: Procedures

Immediately before each of four, regularly scheduled, 40-point multiple-choice examinations, students were told that they could participate in an on-going study to determine how accurately they can predict their own test performance. Students were informed that participation in the study required approximately 5 to 8 minutes of their time and involved completing a short questionnaire. Although

subjects were told that the questionnaire required their name, since their predicted test score and actual test score had to be matched, they were assured that their names and ratings would be kept confidential and that none of their instructors would have access to the data during the semester. To further strengthen this point, an assistant, unknown to any of the students, was in the room to collect the questionnaires immediately after they were completed. After all questionnaires were collected, the assistant left the room before the test was distributed.

Also during the semester, but not on any of the examination days, the TAI, JAS, and BSRI were distributed and completed by students. Once again, students were assured that their identity would be kept confidential from their instructors.

Section Four: Research Design

A multiple-measures, correlational research design was chosen for the study. According to Issac and Michael (1981), the purpose of a correlational design is, "to investigate the extent to which variations in one factor correspond with variations in one or more other factors based on correlation coefficients" (p. 49).

Issac and Michael further suggested that a correlational design is appropriate when variables are complex or, as in the present study, the research does not lend itself to a true experimental design or the controlled manipulation of the independent variables. Correlational designs also allow for the measurement of several variables and their interrelationships simultaneously and in a realistic setting. Finally,

Issac and Michael noted that correlational research gets to the "degree of the relationship" rather than the all-or-nothing question posed by experimental design.

There are, however, certain limitations in selecting a correlational research design (Issac and Michael, 1981). These limitations include the inability of correlational research to identify cause and effect. Correlational research only identifies relationships, which may or may not be causal in origin. A correlational research design is also less rigorous than a true experimental approach because it does not manipulate the independent variables. Correlational research designs may identify spurious relational patterns having little or no reliability and validity. Finally, relational patterns identified in correlational research are often arbitrary and ambiguous. Identifying spurious relationships that have little or no validity and reliability is perhaps the most serious limitation of this study.

As little research has been conducted in the area of understanding how individuals make efficacy judgments, the present study can only be described as exploratory. Consequently, the research concentrates only on determining the types of relationships that exist between accuracy of self-prediction and the stated independent variables. The study will not attempt to establish cause and effect relationships.

Also, due to the nature of the research topic, it is impossible to assign subjects randomly to an accurate prediction group or an inaccurate prediction group. Further, it is not possible to manipulate

the independent variables since for the most part, they represent characteristics (eg., gender, sex-role identity, Type A - Type B behavior continuum, etc.) unique to an individual and cannot be changed or manipulated. Consequently, correlational procedures were deemed the logical and best choice. Also, to facilitate the logistics of a multiple-measure study, intact groups were chosen. By using such groups, it was thought that participant mortality would be minimized.

CHAPTER IV

RESULTS

Statistical results obtained from the study are described in chapter IV. Findings are reported under five major sections. Section one provides an overview of the statistical treatment, including all tested null hypotheses. The second section reports results specific to the Exam Prediction Questionnaire (EPQ). The second section describes the relationship of gender, age, year in school, grade point average (GPA), Jenkins Activity Survey (JAS) score, Test Anxiety Inventory (TAI) score, Bem Sex-Role Inventory score (BSRI), and cumulative test performance to cumulative accuracy score. Section three examines a number of parameters based on JAS scores. The final

Section One: Data Analysis

Statistical data analysis involved three, independent steps. For calculation purposes, all numerical data were treated as either interval or ratio level data. Bivariate and multivariate regression techniques were used in the first step to examine correlations among the various independent variables. The following independent variables were examined for significance and degree of relationship using a Pearson r correlational matrix for each test: (1) predicted test score and actual test score, (2) confidence of prediction and

accuracy score, and (3) predicted test score and confidence of prediction. A multiple regression analysis was conducted to determine the relationship between the independent variables of age, gender, grade point average, year in school, BSRI score, JAS score, TAI score, and actual test performance and the dependent variable of cumulative accuracy score.

In the second independent step, Question #3 of the EPQ in which subjects are asked to rate the degree of influence that each of the information sources exerted while they were formulating a self-prediction, was analyzed. Ratings of the seventeen information sources were treated as independent variables in a multiple regression equation for each of the four tests. The dependent variable was accuracy score.

The third , independent step of the statistical treatment analyzed data collected from Question #4 of the EPQ. Question #4 asks students to choose only the four most important sources of information that they used in making their test prediction. Once these four critical sources were selected, students were further asked to rank them in order of importance from 1 (most important) to 4 (least important). Two statistical treatments were conducted for each of the four tests.

The first treatment was descriptive. Cumulative rankings for each individual information source were determined and forced ranked for each differential score. The second statistical treatment involved using a point biserial r to determine degree of relationship between accuracy score and whether the particular information source was

chosen in the top four or not. According to Isaac and Michael (1981), the point biserial r is chosen when one variable is continuous (accuracy score) and the other represents a genuine dichotomy (in top four or not). Point biserial r's will be calculated for each selected information source on each test.

During the statistical treatment, the following null hypotheses were tested at the .05 alpha level.

1. The correlation between predicted test score and actual test score is 0 for each of the four tests.
2. The correlation between confidence of prediction and accuracy score is 0 for each of the four tests.
3. The correlation between predicted test score and confidence of prediction is 0 for each of the four tests.
4. The correlation between the independent variables of age, gender, grade point average, year in school, BSRI score, TAI score, JAS score and cumulative actual test scores and the dependent variable cumulative accuracy score is 0.
5. The correlation between information cue rating and accuracy score is 0 for each of the four tests.
6. The correlation between accuracy score and critical information cues selected is 0 for each of the four tests.
7. There is no difference in ranking among the information sources for each of the four tests.

Section Two: Findings from the EPQ

Range, mean, and standard deviation for predicted test score,

actual test score, accuracy score, and confidence of prediction for each of the four tests are shown in Table 1. The Pearson correlation coefficient between predicted test score and actual test score, between confidence of prediction and accuracy score, and between predicted test score and confidence of prediction for each of the four tests is displayed in Table 2.

As noted in Table 2, all Pearson correlation coefficients between predicted test score and actual test score are significant at an alpha level of 0.01. None of the Pearson correlation coefficients between confidence of prediction and accuracy score were significant ($p > .05$). Pearson correlation coefficients between predicted test score and confidence of prediction were significant on all four tests ($p < .05$). These findings indicate that at least to a statistically significant degree, university students can predict their own test scores. The findings also indicate that although predicted test score is significantly correlated with confidence of prediction, there is no relationship between confidence of prediction and the accuracy of predicted score.

In part three of the EPQ, students were asked to rate on a 7-point Likert scale the degree of influence each of the seventeen information sources exerted on their test score prediction. Mean ratings for the seventeen information sources for each of the four tests are shown in Table 3. The reader should note that on the first test the variable "performance on past tests in this course" was not calculated since no individual course performance information was available to the students. This practice was followed for all analyses.

A multiple regression equation was calculated for each of the four tests using accuracy score as the dependent variable and the seventeen information sources listed in Table 3 as the independent variables. Of the four calculated equations, only Test *4 was significant ($p < .05$). For Test *1, multiple R was 0.284 and the F value (16, 140) was 0.766 ($p = 0.721$). Multiple R for Test *2 was 0.387 and the F value was (17, 139) was 1.442 ($p = 0.126$). The calculated multiple R for Test *3 was 0.375 and the F value (17, 139) was 1.34 ($p = 0.177$). On Test *4, multiple R was 0.451, squared multiple R was 0.203, adjusted squared multiple R was 0.016, and the F value (17, 139) was 2.08 with a probability of .01. Of the seventeen information sources, only two, self-confidence and interest level, were significantly correlated with accuracy score ($p < .01$). Both information sources were negatively correlated with accuracy score. These results indicate that except in the two cases on Test *4, information source rating concerning perceived influence could not be statistically correlated with accuracy of test prediction.

The seventeen information sources rated in part 3 of the EPQ were subjectively grouped under three general headings for further analysis. The three selected groups involved test-preparation criteria, performance- related criteria, and personal feelings. Information sources subsumed under each heading included: Test-preparation criteria: (1) amount of effort exerted in studying for the test, (2) knowledge of the material to be covered on the test, (3) how well the teacher presented the material to be covered on the test, (4) concentration level while studying for the test, (5) how

interesting the material covered by the test was for the student, and (6) the perceived effectiveness of the student's study skills for the test.

Performance-related criteria: (1) general academic ability, (2) performance in other educational courses, (3) performance on previous tests in the course, (4) perceived test difficulty, and (5) the performance of other students in the course.

Personal feelings: (1) self-confidence, (2) mood, (3) physical health, (4) anxiety level, (5) amount of encouragement given by the teacher, and (6) luck.

Mean ratings of the three groups for each test were calculated and are shown in Table 4. As indicated in Table 4, test-preparation criteria was rated the highest followed by performance-related criteria and personal feelings. A one-way, within subjects analysis of variance was calculated for each test to determine if significant differences existed among test-preparation criteria, performance-related criteria, and personal feelings. No significant differences were detected, indicating that students did not preferentially rate one group of information cues higher than the other two groups.

In part four of the EPQ, students were asked to select the four most important sources of information that they used in making their test predictions. Once the four information sources had been selected, the students were further asked on the questionnaire to rank them from 1 (most important) to 4 (least important). Cumulative ratings for each information source were calculated using a scoring system

of 4 for an information source given a 1 in part 4 of the EPQ, a 3 for an information source given a 2, a 2 for an information source given a 3, a 1 for an information source given a 4, and a 0 for an information source not selected. All cumulative ratings for the information sources for each test are summarized in Table 5. In Table 6, the information sources, based on their cumulative ratings, are ranked in order of importance from the most important (1) to the least important (17).

The seventeen information sources were also grouped under the three general headings of test-preparation criteria, performance-related criteria, and personal feelings. The individual information sources comprising the three groups are the same as previously described. Mean ratings of the three groups were calculated and are shown in Table 7. As indicated in Table 7, test-preparation criteria was rated the highest, followed by performance-related criteria and personal feelings.

A one-way within-subjects analysis of variance was calculated for each of the four tests to determine if significant differences existed among the three groups. If significant differences were detected, it would support the hypothesis that in a forced choice situation, students would select information cues from one group preferentially over the other two groups. All analyses were significant at an alpha level of .05. For Test #1, the F value was 7.5 ($p = 0.023$), for Test #2 F was 11.69 ($p = .004$), for Test #3 F was 7.44 ($p = .015$), and for Test #4 the F value was 6.83 ($p = .019$). A follow-up Tukey's HSD post-hoc test indicated significant differences

($p < .05$) between test-preparation criteria and personal feelings for Test #1, #2, #3, and #4. Also, significant differences were detected between test-preparation criteria and performance-related criteria for Test #2, #3, and #4. No significant differences were indicated, however, between performance-related criteria and personal feelings for any of the four tests. Thus, results suggest that when students were in a forced choice situation, they preferentially selected test-preparation cues over performance-related and personal feeling cues.

Each individual information source in part four of the EPQ was further correlated with accuracy score for each of the four tests using a point biserial Pearson correlation coefficient. For computational purposes, a 1 was assigned if the information source was one of the four selected, irrespective of ranking, and a 0 if the source was not selected. Results of all correlation calculations are shown in Table 8. As indicated in Table 8, no significant correlations ($p > .05$) were observed, indicating no relationship between cue selection and accuracy of predicted test score.

Section Three: Multiple Regression Analysis

A multiple regression equation was calculated to explore the relationship of the independent variables of age, gender, year in school, university grade point average, JAS score, TAI score, BSRI score, and cumulative actual test performance to the dependent variable cumulative accuracy score. Cumulative accuracy scores represent the sum of the four, individual accuracy scores.

Results from the calculated regression equation are summarized in Table 9. As reported in Table 9, multiple R was 0.756, squared multiple R was 0.572, and the adjusted squared multiple R was 0.548. Multiple R was significant ($p < .01$) with an F value (8,148) of 24.68. Also, as noted in Table 9, only two independent variables, JAS score and cumulative test performance, were significant at an alpha level of .01. GPA was significant at an alpha level of 0.05. Further, JAS score was negatively correlated to the dependent variable whereas cumulative test performance and GPA were positively correlated.

The squared semi-partial of each independent variable was also calculated. Based on these calculations, age accounted for 0.29% of the variability in cumulative accuracy score, gender 0.32%, year in school 0.01%, grade point average 1.46%, JAS score 13.06%, TAI score 0.96%, BSRI score 0.19%, and cumulative test performance 10.97%. These results indicate that JAS score and cumulative test performance accounted for approximately 24% of the unique variance in cumulative accuracy score while the other six independent variables accounted for only approximately 3.2%. These findings suggest that at least in this study, actual student performance and the Type A - Type B behavior pattern continuum had a significant impact on the accuracy of test prediction.

Section Four: JAS Score Characterization

As previously described, JAS scores were found to be significantly related to cumulative accuracy scores. In an attempt to

further investigate the influence of Type A - Type B behavior pattern on predictive accuracy, JAS scores were trichotimized into three groups representing the Type A behavior pattern (JAS score > than 10), the Type AB behavior pattern (JAS score between 6 and 10), and the Type B behavior pattern (JAS score of less than 5). Similar divisions have been used by Ward and Eisler (1987) and Krantz, Glass, and Snyder (1974). Based on these divisions, 39 (25%) individuals were identified as exhibiting a Type B behavior pattern, 55 (35%) individuals as exhibiting a Type AB behavior pattern, and 63 (40%) individuals as exhibiting a Type A behavior pattern.

Pearson correlation coefficients between predicted test score and actual test score were recalculated for each behavior pattern on each test. All recalculated Pearson correlation coefficients were significant ($p < .01$) and are illustrated graphically in Figure 1. As noted in Figure 1, Type B individuals had the highest correlation coefficients on all four tests. A formula described by Cohen and Cohen (1983, p.54) for testing the significance of the difference between Pearson correlation coefficients obtained on two independent r's was used to compare Type A - AB - B behavior patterns. Specifically, predicted test score and actual test score correlation coefficients were compared for each test among the three behavior patterns.

For Test #1, a significant difference was detected for the correlation coefficients between Type AB and Type B behavior patterns ($z = -1.98$, $p < .05$). No significant differences were noted, however, between Type A and Type B behavior patterns ($z = -1.69$, $p > .05$) or between Type A and Type AB behavior patterns ($z = 0.38$, $p >$

.05).

On Test *2, a significant difference was detected for the correlation coefficients between Type A and Type B behavior patterns ($z = -2.189$, $p < .05$. No significant differences were found between Type AB and Type A behavior patterns ($z = -1.50$, $p > .05$) or between Type A and Type AB behavior patterns ($z = -0.72$, $p > .05$).

For Test *3, significant differences were noted for the correlation coefficients between Type A and Type B behavior patterns ($z = -2.89$, $p < .01$) and between Type AB and Type B behavior patterns ($z = -2.30$, $p < .05$). No significant difference was detected between Type A and Type AB behavior patterns ($z = -0.30$, $p > .05$).

Significant differences for Test *4 were observed for the correlation coefficients between Type A and Type B behavior patterns ($z = -2.58$, $p < .05$) and between Type AB and Type B behavior patterns ($z = -2.53$, $p < .05$). No significant difference was detected between Type A and Type AB behavior patterns ($z = 0.01$, $p > .05$). The four test comparisons indicate that Type B individuals significantly predicted subsequent test performance more accurately than their Type A and Type AB counterparts.

Cumulative accuracy scores for the four tests were subjectively divided into three groups: < -4, -4 to $+4$, and $> +4$. Negative accuracy scores indicate that individuals overestimated their actual test performance while positive accuracy scores indicate that individuals underestimated their actual test performance.

Sixty-nine individuals (44%) had cumulative accuracy scores of less than -4. Of those 69 individuals, 53 or 77% possessed Type A

behavior patterns, 13 or 19% displayed Type AB behavior patterns, and 3 or 4% possessed Type B behavior patterns. Fifty-eight individuals (37%) had cumulative accuracy scores of between -4 and +4. Of those 58 individuals, 6 or 10% possessed Type A behavior patterns, 22 or 38% displayed Type AB behavior patterns, and 30 or 52% displayed Type B behavior patterns. Thirty individuals (19%) had cumulative accuracy scores greater than +4. Of these 30 individuals, 4 or 13% possessed Type A behavior patterns, 20 or 67% possessed Type AB behavior patterns, and 6 or 20% possessed Type B behavior patterns. These results are presented graphically in Figure 2.

The findings were also entered into a 3 x 3 Chi square analysis using behavior type (A, AB, and B) and cumulative accuracy scores (< -4, -4 to +4, and > +4) as the matrix headings. Calculated Chi square was 109.02 with eight degrees of freedom. The calculated Chi square was significant ($p = .0001$).

Information sources selected in part four of the EPQ were reexamined to detect differences among Type A - AB - B behavior patterns. The information sources were divided into the three groups previously described: test-preparation criteria, performance-related criteria, and personal feelings. A between subjects, one-way analysis of variance was calculated to determine if significant differences existed among the three behavior patterns for each major information group. None of the F values were significant at an alpha level of .05. For Test #1, the F value for test-preparation criteria was .02, for performance-related criteria .02, and for personal feelings 0.19. For Test #2, the F value for test- preparation criteria was .137, for

performance-related criteria .23, and for personal feelings 0.38. For Test #3, the F value for test-preparation criteria was .052, for performance-related criteria .052, and for personal feelings 0.10. For Test #4, the F value for test-preparation criteria was .05, for performance-related criteria .03, and for personal feelings 0.05.

Section Five: Summary of Statistical Findings

Findings from the various statistical treatments are summarized accordingly:

1. The Pearson correlation coefficient between predicted test score and actual test score was significant ($p < .01$) for each of the four tests. These findings reject the null hypothesis stating that the correlation between predicted test score and actual test score is 0. The study found that at least to a statistically significant degree, students can predict their own test performance.

2. The Pearson correlation coefficient between confidence of prediction and accuracy score was not significant ($p > .05$) for any of the four tests. These findings fail to reject the null hypothesis stating that the correlation between confidence of prediction and accuracy score is 0 for each of the four trials. Apparently, confidence in making a test score prediction is not related to the accuracy of that prediction.

3. The Pearson correlation coefficient between predicted test score and confidence of prediction was significant ($p < .01$) for each of the four tests. These findings reject the null hypothesis stating that the correlation between predicted test score and confidence of prediction

is 0 for each of the four tests. Higher predicted test scores were significantly correlated with higher confidence levels.

4. In general, none of the seventeen information sources were significantly correlated ($p > .05$) with accuracy score on any of the four tests. The only exception to this generalization was on the final test (Test #4), when self-confidence and interest level were significantly correlated ($p < .05$) with accuracy score. Both information sources were negatively correlated. These findings generally fail to reject the null hypotheses stating that the correlation between information cue rating and accuracy score is 0 for each of the four tests and that the correlation between accuracy score and critical information cue selected is 0 for each of the four tests. The study failed to indicate a significant relationship between efficacy cue selection and/or weighting and accuracy of test prediction.

5. The independent variables age, gender, year in school, GPA, JAS score, TAI score, BSRI score, and cumulative test performance were significantly correlated ($p < .01$) with cumulative accuracy score. These findings reject the null hypothesis stating that the correlation between the independent variables of age, gender, grade point average, year in school, BSRI score, TAI score, JAS score, and cumulative actual test scores and the dependent variable cumulative accuracy score is 0. Of the eight entered independent variables, only GPA, JAS score, and cumulative test performance were significant ($p < .05$). GPA and cumulative test performance were positively correlated with cumulative accuracy score, whereas JAS score was

negatively correlated.

6. When JAS scores were trichotimized into Type A (>10), Type AB (6 - 10), and Type B (< 6) behavior patterns, recalculated Pearson correlation coefficients between predicted test score and actual test score indicated significant correlations ($p < .01$) for all three groups on each test. However, Type B individuals had significantly higher correlations than either Type A or Type AB individuals. Of those individuals overestimating their test performance, 77% were identified as Type A, 19% as Type AB, and 4% as Type B. Conversely, of those individuals underestimating their test performance, 13% were identified as Type A, 67% as Type AB, and 20% as Type B. Of those individuals accurately predicting test performance (a cumulative accuracy score of plus/minus 4), 52% were identified as Type B, 10% as Type A, and 38% as Type AB.

7. When information sources were grouped under test-preparation criteria, performance-related criteria, and personal feelings, no significant differences were detected when rankings from part 3 of the EPQ were calculated. However, significant differences were noted among the three efficacy cue groups for each test when students had to make a forced choice in part 4 of the EPQ. Test-preparation criteria was preferentially selected over performance- related criteria and personal feelings. No significant differences were detected, however, between performance-related criteria and personal feelings.

CHAPTER V

DISCUSSION

Chapter V discusses results obtained from the study and is divided into six sections. Section one discusses the relationships among predicted test score, actual test score, and confidence of prediction. Section two discusses results pertaining to the first underlying question of the study which asks whether there are significant differences in the types of information sources selected and/or weighted among students who make accurate predictions of test scores when compared to those who do not. The second underlying question of the study is discussed in Section three and pertains to whether age, gender, GPA, year in school, BSRI score, JAS score, TAI score, and cumulative actual test score can be correlated with the accuracy of predicted test performance. Observations made in Sections one, two, and three are synthesized in Section four in order to present an integrated theoretical model concerning the accuracy of efficacy judgments. Section five offers suggestions for additional research concerning the identification of factors affecting the accuracy of efficacy judgments. Section six concludes the chapter by summarizing the major findings from the study.

Before beginning a discussion of the results, a few words of caution may be appropriate. First, it must be remembered that

correlational studies do not prove a causal relationship. Second, a fairly homogeneous, intact group was used throughout the study. In using such a group in a correlational study, there is always the chance of a restricted range. Also, the extent that the findings may be generalized to a more heterogeneous group is unknown. Third, the instruments used and the course tests taken be the students possessed varying reliabilities. Attenuation of statistical results may be anb inherent danger in the study. Finally, the semi-partials for the multiple regression equation should be interpreted as representing only relative, not absolute contributions.

Section One: Test Predictability

Results from the study indicate that, at least to a statistically significant degree, university students can predict their own test performance. Pearson correlation coefficients between predicted test score and actual test score stayed fairly constant throughout the four tests (.50 +/- .03), with the highest correlation coefficient (0.531) occurring on the first test. This result is somewhat surprising since no specific course performance information was available to students during their first test. It was originally anticipated that students would use their first test performance as a "benchmark" for subsequent test predictions and that such predictions would become more accurate as the course progressed. This anticipated improvement, however, was not observed.

A suggested explanation for this overall lack of improvement in predicted test accuracy is that students may not fully integrate past

performance attainments into future performance predictions. Instead of focusing on how well they had done on previous tests, using this information as a performance base rate, students may have instead focused on information cues dealing with immediate test preparation. This same judgmental bias was observed earlier by Kahneman and Tversky (1973) and termed "representativeness." As will be discussed more fully in Section two of this chapter, there is strong evidence that the judgmental bias of representativeness occurred during the present study.

Although confidence of prediction was significantly correlated with predicted test score, it was not correlated with accuracy score. These results suggest that students making higher test predictions were more confident in their predictions than students making lower test predictions. These elevated confidence levels, however, were not related to the accuracy of test predictions. In many instances, high predicted test scores accompanied by high confidence levels represented an overestimation of both performance potential and appropriate confidence level.

In earlier studies concerning confidence level, Bandura and Cervone (1983) demonstrated that self-confidence positively affected task effort and task choice. Higher confidence levels for example, were positively correlated with higher task effort. Although confidence level and task effort and choice may be significantly correlated, the present study indicates that confidence level has little impact on the accuracy of efficacy judgments. In fact, overconfidence may negatively affect the accuracy of self-

predictions.

Two conclusions may be drawn from this portion of the study. First, at least in a limited sense, students can predict their own test scores. Prediction accuracy, however, did not improve with each subsequent test score. Instead, accuracy correlations stayed fairly constant over the four test period. This lack of improvement suggests that students may not have focused on information cues relating to actual test performance as much as they could have. Similar findings were made by Kahneman and Tversky (1973) who noted that individuals commonly fail to integrate past performance attainments into future performance predictions even when, as in the present study, this information is readily available.

The second conclusion drawn from the study is that student self-confidence level, as stated on the EPQ, does not impact the accuracy of efficacy judgments. Although Bandura and Cervone (1983) found confidence level to positively influence task effort and task choice, the present study found no relationship between self-confidence and accuracy of self- prediction. These findings suggest that confidence level is not a reliable indicator of actual performance attainments.

Section Two: Efficacy Information Sources

The study attempted to determine whether there are significant differences in the types of efficacy information sources selected and/or weighted among students who make accurate predictions of test performance when compared to those who do not.

Based on the present study's design and instrumentation, it is concluded that at least for the studied group, no such differences existed. A significant correlation between accuracy score and efficacy information source was observed on only one test (Test #4). In this particular case, the correlated information sources were self-confidence and interest level. Both information sources were negatively correlated with accuracy score.

The finding concerning self-confidence suggests that overconfidence can lead individuals to overestimate performance potential, while individuals experiencing low levels of self-confidence are more apt to underestimate subsequent performance potential. This same observation may also apply to interest. Students highly interested in a subject may feel that they will do well on test-related material, thereby inflating their predicted performance potential. Conversely, students not interested in a subject may feel that they will not perform well, thereby lowering their judgments of test performance.

Thus, findings concerning efficacy information source selection and weighting suggest that in the present study, students, regardless of the accuracy of their test predictions, generally selected and weighted the various efficacy information sources in a similar manner. This selection and weighting process did not substantially differentiate those students making accurate test predictions from those greatly overestimating or underestimating subsequent test performance.

Although the study failed to identify a correlation between

efficacy cue selection and/or weighting and accuracy of prediction, it did reveal that all students, regardless of their ability to accurately predict their own test performance, consistently selected certain efficacy cues preferentially over others. Specifically, the study found that most students generally selected and weighted efficacy information cues dealing with test-preparation criteria preferentially over either performance-related criteria or personal feelings. Test-related knowledge, effort devoted to test preparation, and concentration level while studying were the top three choices of students on all four tests on both parts three and four of the EPQ. This overwhelming focus on test preparation cues at the expense of performance-related cues may explain why predicted test score and actual test score correlation coefficients did not improve with each subsequent test. This concept will be discussed further in Section four of this chapter.

Concerning performance-related cues, the highest rating for performance-related criteria (Table 7) was obtained on the first test, which also had the highest correlation coefficient (0.531) between predicted test score and actual test score as discussed in Section one. A 29% drop in the mean rating of performance-related criteria, however, occurred from the first test to the second. After the first test, students did not seem to place as much importance on performance-related criteria. This observation is in agreement with the earlier findings of Kahneman and Tversky (1973) who noted that individuals habitually fail to integrate actual past performance attainments into future performance predictions, even when this

information is readily available. Apparently, both Kahneman and Tversky's subjects and the students participating in this study placed greater importance on cues associated with subsequent task preparation than previous task accomplishments. This judgmental bias commonly exists despite repeated studies indicating that more accurate performance predictions can be made when actual past performance accomplishments are integrated into the predictive process (Arkes and Hammond, 1986).

Findings from the present study also question the suggestion of Bandura (1982) that efficacy judgments are primarily based on actual past performance accomplishments, observing the performance of others, verbal persuasion or coaching, and physiological (anxiety) state. When the 17 efficacy information sources were forced ranked (Table 6) from a high of one to a low of 17, actual classroom performance had a mean ranking of six and performance in other courses a mean ranking of 13. Vicarious observations of other students' performance received a mean ranking of 16, anxiety level a mean ranking of 11, and verbal persuasion (encouragement given by the teacher) a mean ranking of 17. Throughout the study, students apparently did not place a great deal of emphasis on actual performance attainments, whether their own or those of other students, while making predictions of test performance. Also, verbal encouragement given by the teacher and perceptions of anxiety were rated quite low. At least in the present study, Bandura's major efficacy sources were not perceived by students as being as important as previously thought.

Conversely, the major achievement attributes of effort, ability, perceived task difficulty, and luck proposed in attribution theory (Weiner, 1985) generally received much higher rankings. Effort received a mean ranking of two, general academic ability a mean ranking of nine, knowledge of the material (a reflection of ability) a mean ranking of one, perceived test difficulty a mean ranking of five, and luck a mean ranking of 15.

These results suggest that students in the present study placed a much greater emphasis on the amount of effort spent preparing for a test and the knowledge gained from that effort than the actual achieved test scores themselves. A possible explanation for this phenomena is that students correlated amount of effort expended and knowledge gained to specific test scores. In making subsequent test score predictions, students would then compare their current effort expenditure and achieved knowledge to past test-related perceptions of effort and knowledge. For example, if one unit of effort resulted in a test score of 30, and the student currently feels he or she expended two units of effort in preparing for the next test, a predicted test score may be adjusted accordingly (for example increased to 35) to reflect this increased effort. Therefore, the primary focus is on the amount of effort devoted to test preparation which in turn is a representation of past performance. Thus, actual past performance attainments may be integrated into the formulation of future performance predictions only indirectly through the determination of causal attributes such as effort, knowledge, or concentration level.

Three major conclusions are drawn from this portion of the

study. First, efficacy information cue selection and weighting did not significantly differentiate those students making accurate test predictions from those students who did not. Students who under-estimated, overestimated, and accurately predicted test performance generally selected the same information sources and weighted them similarly.

The second finding indicates that students, regardless of accuracy scores, chose efficacy information cues dealing with test preparation preferentially over those dealing with either past performance attainments or personal feelings, such as mood or health. Although Kahneman and Tversky (1973) argue that integrating and utilizing past performance data allows for more accurate intuitive predictions, their studies repeatedly demonstrated that individuals rarely utilize this data source, even when readily available. Similar conclusions to those of Kahneman and Tversky are drawn from this study.

The third finding indicated that the top six selected efficacy information sources in descending order of perceived importance were test-related knowledge, effort spent during test preparation, concentration level while studying, effectiveness of study skills while preparing for a test, perceived test difficulty, and actual classroom performance on past tests. The study suggests that achievement attributes identified in attribution theory seem to play a more important role in formulating efficacy judgments than those originally proposed by Bandura (1982). These findings appear to support Schunk's (1984) hypothesis that an individual's attributions

concerning past performance or failure influence subsequent efficacy judgments. It is suggested that individuals may cognitively correlate actual performance attainments with specific attributes such as effort, concentration level, or achieved knowledge, and it is these attributes that are are selected and weighted during the formulation of efficacy judgments and not the actual performance attainments themselves. Therefore, Bandura's efficacy sources may not play as great a role in efficacy judgment formulation as previously thought.

Section Three: Individual Differences

Statistical results from the study indicated that approximately 54% of the variance in cumulative accuracy score could be accounted for by the independent variables of age, gender, year in school, GPA, JAS score, TAI score, BSRI score, and cumulative actual test scores (Table 9). Of the eight variables, however, only JAS score, actual test performance, and GPA were significantly correlated with cumulative accuracy score. The earlier hypotheses developed in Chapter II that anxiety level and sex-role identity would affect the accuracy of test predictions were not supported.

Two possible explanations for the lack of a significant correlation between anxiety level as measured by the TAI and cumulative accuracy score can be offered. First, Bandura's (1986) original hypothesis that elevated levels of anxiety affects the accuracy of efficacy judgments may not be valid or may be too simplistic. Although undue anxiety may negatively affect actual performance attainments, elevated anxiety levels may not adversely

affect predictions of performance. In such instances, actual performance and expected performance may be quite different. A second explanation for the lack of a significant correlation is that the TAI measures trait anxiety and not state anxiety. State anxiety levels would most likely fluctuate more than trait anxiety levels. A measure of state anxiety immediately before each test, therefore, may have revealed a significant correlation between accuracy of prediction and anxiety level. However, in the present study, such measurements were not taken due to the realistic constraint of allowable time before each test. It simply wasn't possible in the current study to submit students to a lengthy battery of instruments immediately before an important examination. For this reason, the TAI was chosen since it could be given to all students on a non-test day.

Cumulative accuracy scores and sex-role identity BSRI scores were also not correlated. At least in this study, perceived sex-role had no impact on the ability to accurately predict test scores. Apparently, perceived sex-role identity, whether highly masculine, highly feminine, or androgenous, does not affect the accuracy of efficacy judgments.

The calculated multiple regression did reveal, however, that grade point average and actual test scores were positively correlated with cumulative accuracy score. The study indicated that lower performing students frequently overestimated their own test performance potential. Conversely, students of much higher abilities frequently underestimated their performance potential. The most accurate test predictors were slightly above average students, those

receiving an equivalent letter grade of C+ or B- on the four tests.

Lower performing student consistently predicted higher test scores than actually achieved. These students apparently did not place a great deal of importance on repeated poor performances. Their continued focus on test preparation cues at the expense of performance cues may account for their inability to accurately predict test performance. High ability students frequently underestimated their own performance potential. Apparently, these students do not believe that they can consistently perform as well as they actually do. In contrast, average students predicted their own test performance fairly accurately during the study. It appears that most students of average ability possess a realistic view of their own performance capabilities. Yet as noted in the previous section, these students could not be differentiated based on information cue selection or weighting.

The findings concerning actual test performance indicate that ability, as reflected by actual test scores, is significantly correlated with accuracy of prediction. Lower performing students appeared to possess an inflated opinion of their own abilities, while higher performing students apparently did not trust their repeated successes. Average students appeared to have a much more realistic view of their own performance capabilities.

The Type A - Type B behavior pattern was also shown to be significantly correlated with cumulative accuracy score. Statistical analysis revealed a negative correlation, indicating that Type A individuals consistently overestimated their actual test scores,

while Type AB and B individuals tended to underestimate or more accurately predict test performance. This finding concurs with earlier research by Grimm and Yarnold (1984) and Ward and Eisler (1987) whom also found Type A individuals repeatedly overestimating their performance potential. The present study also indicated that Type B individuals were frequently more accurate predictors of their own test performance than their Type AB counterparts. Throughout the study, Type AB individuals frequently underestimated their own test performance, although not to the degree that Type A individuals overestimated their test scores. This distinction between the Type AB and Type B behavior pattern has not been previously reported in the literature. In most previous studies, the focus of the research has been primarily on the inability of the Type A individual to meet predicted performance goals. Consequently, Type AB and Type B individuals have frequently been grouped together as individuals who either underestimated or fairly accurately predicted their own performance capabilities (Ward and Eisler, 1987). By treating the Type A and Type AB behavior patterns individually in the present study, a finer distinction could be drawn.

The study, however, revealed no significant differences in the efficacy information cues selected by Type A, Type AB, and Type B individuals. All behavior patterns selected and weighted efficacy information sources in a similar manner. This finding indicates that at least concerning the present study, accuracy of test prediction and the Type A - Type B behavior continuum could not be differentiated based on efficacy information source selection and/or weighting.

To summarize, the Type A - Type B behavior continuum appears to significantly affect the accuracy of efficacy judgments. Specifically, the study revealed that the Type A behavior pattern repeatedly overestimated test performance, the Type AB behavior pattern frequently underestimated test performance, and the Type B behavior pattern was the most accurate predictor of test performance. No significant differences were noted, however, in the efficacy cues selected and/or weighted by the three behavior patterns.

Section Four: Synthesis

The present study suggests that the formulation and accuracy of efficacy judgments involves a number of interrelated variables. For example, factors that may account for individuals grossly overestimating their performance potential include a strong Type A behavior pattern, relatively low ability, overconfidence, and a preferential concentration on task preparation cues at the expense of performance- related cues. Conversely, individuals accurately estimating their own performance potential would most likely be characterized by a Type B behavior pattern, average ability, and an appropriate confidence level.

Although these various factors can "profile" the accurate predictor of self-performance from someone who either overestimates or underestimates performance potential, they fail to adequately explain the "why" of how efficacy judgments are actually made. It is suggested that the answer still lies in how efficacy cues are selected and weighted. Although the present study failed to

identify a linkage between accuracy of efficacy judgment and information source selection and weighting, it is suggested that additional research and instrument redesign may discern such a connection.

One possible explanation of the "why" is to consider Type A individuals who consistently overestimate their own performance potential. Recall that Wright (1988) suggested that Type A individuals are exposed to timed activities quite early in their life. Such exposure argues Wright, provides a personal blueprint for achieving more by effectively managing time and by chronic activation. The Type A individual, however, may transform this observation into a somewhat different meaning. Specifically, the Type A individual may learn that increased effort, a representation of chronic activation, leads to increased performance attainments. Such an early association between effort and performance may limit the Type A individual from focusing on other causal performance attributions. In predicting task performance for example, the Type A individual would correlate effort expenditure to level of task accomplishment. Accordingly, a high effort expenditure should result in a high level of task accomplishment.

There is empirical research to support this effort/performance link. Terborg (1977) for example, measured the impact of human effort and ability on two tasks of differing difficulty. He observed that effort was a significant factor in accounting for performance variance on both the simple and difficult task. Ability, however, was statistically significant only on the difficult task.

If we assume that the Type A individual had past experiences with, and was reinforced through, the correlation between effort and performance at an early age while dealing with relatively simple tasks, it may be possible that this perceived correlation carried over into all tasks, irrespective of task difficulty. Such an individual would then associate all subsequent efficacy judgments with effort expenditure while ignoring both ability and actual past performance attainments. Poor performances could be justified through a lack of effort and an increased determination to "try even harder" next time. However, as noted by Terborg (1973), effort is only one important variable affecting performance on difficult tasks: the other is ability. Yet the Type A individual may fail to make this connection, instead always relying on a personal assessment of effort. In the current study, many information cues may be intuitively correlated to effort. For example, increased effort may be intuitively correlated to knowledge. In such thinking, high effort expenditure must mean high knowledge gain.

If other factors such as self-esteem and actual ability are considered, it is easy to view the formulation of efficacy judgments as a highly complex cognitive process. For example, consider the interactive dynamics of an individual characterized as having a Type B behavior pattern, high self-esteem, moderate ability, and a primary focus on past task performance, in contrast to someone with a Type A personality, low ability, low self-esteem, and a preoccupation with task effort. It is suggested that such differential characteristics would affect the way efficacy judgments are formulated.

The current study also revealed an overwhelming preoccupation of students with task preparation cues at the expense of past performance cues. Apparently, students did not process actual performance attainments but rather the causal attributes associated with those attainments. Such cognitive correlations appear to form the basis for subsequent efficacy judgments. For example, an actual performance attainment is associated with a specific causal cue(s), such as effort or ability. Not only are such cues identified, but they also appear to be quantified in such a way that subsequent efficacy judgments are based on this quantification. Increasing or decreasing the amount of effort on subsequent tasks for example, would hypothetically result in a corresponding lowering or raising of efficacy judgments as well. If this hypothesis is correct, then specific personality characteristics such as the Type A - Type B behavior continuum may simply represent a reflection of this correlational process. For example, Type A individuals may intuitively associate expended effort with actual performance attainments to a greater degree than do their Type B counterparts. Obviously, more research is needed to validate this hypothesis.

Section Five: Future Research Directions

The following suggestions are made for additional research concerning the identification of factors affecting the accuracy of efficacy judgments.

1. No attempt was made in the present study to quantify the various efficacy information sources. It is suggested that future studies

ascertain the individual's quantitative perceptions of each causal information source (e.g., high, moderate, or low perceived effort) and determine if subsequent efficacy judgments are related to fluctuations in this weighting process.

2. The study also did not investigate the impact of state anxiety on the accuracy of efficacy judgments. If state anxiety measurements had been taken in the present study, state anxiety levels may have supported Bandura's (1986) hypothesis that individuals suffering elevated anxiety would underestimate performance potential. One realistic constraint in the present study concerning both state anxiety and information cue quantification was time. It was simply not possible to give lengthy instruments to students immediately before taking a major university examination. Although these measurements could be given in a controlled research setting where time constraints are not of critical importance to the study participants, the validity of such measurements may be questioned. For example, would an individual experience the same anxiety level in a controlled setting as opposed to a real-life situation, such as a classroom testing environment?

3. The role of the Type A - Type B behavior pattern in affecting the accuracy of self-predictions appears especially intriguing and in need of additional research. Two different lines of inquiry may prove informative. First, studies should be developed to further investigate the overall estimating abilities of the Type A versus Type B individual. Are their significant differences in the way the two personality types generally make estimates of performance, whether

dealing with human performance or some other type of performance indicator? A second line of inquiry should investigate the correlation between the accuracy of efficacy judgments and the independent variables of risk taking, self-esteem, and the Type A - Type B behavior pattern. Are Type A individuals greater risk takers than their Type B counterparts and is this difference significantly correlated with the accuracy of self-predictions? Also, does the Type A individual, as suggested by Wright (1988), commonly possess lower self-esteem and if this is the case, do these self-perceptions affect the ability to accurately estimate one's own performance capabilities? Both lines of inquiry may prove beneficial in better understanding the predictive accuracy of the Type A individual when compared to the Type B individual.

4. A final avenue of inquiry may begin to explore how to develop more realistic perceptions of performance attainments within individuals. Specifically, how do we improve the accuracy of efficacy judgments?

Section Five: Summary

Results from the study indicated that at least within a limited sense, students can predict their own test performance. Pearson correlation coefficients between predicted test score and actual test score did not improve, however, over the four test period. Apparently, students did not incorporate past performance attainments (actual test scores) into future test predictions. These findings support earlier work by Kahneman and Tversky (1973) concerning the judgmental bias of representativeness.

Although student confidence level concerning test score predictions was significantly correlated with predicted test score, no significant correlations were noted between confidence of prediction and accuracy score. It appears that confidence level does not affect the accuracy of efficacy judgments.

Significant differences between efficacy information cue selection and/or weighting were observed on only one test. Throughout the study, students generally selected and weighted efficacy information cues similarly irrespective of test prediction accuracy. However, differences were noted among the importance students placed on various information sources, irrespective of predictive accuracy. Generally, those information cues associated with test preparation were preferentially chosen or selectively weighted over those cues dealing with either past performance or personal feelings. These findings question the importance of efficacy cues proposed by Bandura (1982) and better support achievement attributes identified by Weiner (1979). Apparently, efficacy judgments rely heavily on causal ascriptions.

Actual performance attainments were significantly correlated with cumulative accuracy score. High performing individuals frequently underestimated their performance potential, while low achieving students frequently overestimated their test performance. In the present study, average students were the most accurate predictors of test performance.

The Type A - Type B behavior continuum was also significantly correlated with cummulative accuracy score. Type A individuals

repeatedly overestimated their test performance. These observations support similar findings by Grimm and Yarnold (1984) and Ward and Eisler (1987). The study also found that Type AB personalities were more apt to underestimate their test performance while Type B personalities were the most accurate predictors of test scores.

Although the present study identified specific factors affecting the accuracy of efficacy judgments, it failed to understand how efficacy information is cognitively processed. Self-reports of efficacy information cue selection and weighting did not reveal significant differences concerning the accuracy of efficacy judgments. Future studies concerning the accuracy of efficacy judgments should concentrate on developing a better understanding of how indivduals process mutidimensional information. Specifically, it is suggested that future studies ascertain the individual's quantitative perceptions of each causal information source (e.g., high, moderate, or low perceived effort) and determine if subsequent efficacy judgments are related to fluctuations in this weighting process.

REFERENCES

Anderson, C.A. (1983). The causal structure of situations: The generation of plausible causal attributions as a function of type of event situation. Journal of Experimental Social Psychology, 19, 185-203.

Arkes, H.R., and Hammond, K.R. (Eds.). (1986). Judgment and decision making: An interdisplinary reader. Cambridge: Cambridge University Press.

Bakan, D. (1966). The duality of human existence. Chicago: Rand McNally, 1966.

Bandura, A. (1977). Social learning theory. Englewood Cliffs, New Jersey: Prentice Hall.

Bandura, A. (1981). Self-referent thought: A developmental analysis of self-efficacy. In J.H. Flavell & L.D. Ross (Eds.), Cognitive social development: Frontiers and possible futures (pp. 200-239). New York: Cambridge Press University.

Bandura, A. (1982). Self-efficacy mechanism in human agency. American Psychologist, 37, 122-147.

Bandura, A. (1986) Social foundations of thought & action. Englewood Cliffs, New Jersey: Prentice Hall.

Bandura, A., Adams, N.E., & Beyer, J. (1977). Cognitive processes mediating behavioral change. Journal of Personality and Social Psychology, 35, 125-139.

Bandura, A., Adams, N.E., Hardy, A.B., & Howells, G.N. (1980). Tests of the generality of self-efficacy theory. Cognitive Therapy and Research, 4, 39-66.

Bandura, A. and Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. Journal of Personality and Social Psychology, 45, 1017-1028.

Barling, J., and Beattie, R. (1983). Self-efficacy beliefs and sales performance. Journal of Organizational Behavior Management, 5, 41-51.

Bar-Tal, D., Goldberg, M., and Knaani, A. (1984). Causes of success and failure and their dimensions as a function of SES and gender. A phenomenological analysis. British Journal of Educational Psychology, 54, 51-61.

Bem, S.L. (1974). The measurement of psychological androgyny. Journal of Consulting and Clinical Psychology, 42, 155-162.

Bem, S.L. (1981). Bem Sex-Role Inventory. Palo Alto, CA: Consulting Psychologists Press, Inc.

Betz, N.E. and Hackett, G. (1981). The relationship of career-related self-efficacy expectations to perceived career options in college women and men. Journal of Counseling Psychology, 23, 399-410.

Biran, M., and Wilson, G.T. (1981). Treatment of phobic disorders using cognitive and experience methods: A self-effficacy analysis. Journal of Counseling and Clinical Psychology, 49, 886-889.

Brown, I., Jr., and Inouye, D.K. (1978). Learned helplessness through modeling: The role of perceived similarity in competence.

Journal of Personality and Social Psychology, 36, 900-908.

Burger, J.M., Cooper, H.M., and Good, T.L. (1982). Teacher attributions of student performance: Effects of outcome. Personality and Social Psychology Bulletin, 4, 685-690.

Campbell, N.K. and Hackett, G. (1986). The effects of mathematics task performance on math self-efficacy and task interest: Journal of Vocational Behavior, 28, 149-162.

Chambliss, C.A., & Murray, E.J. (1979) Cognitive procedures for smoking reduction: Symptom attribution versus efficacy attribution. Cognitive Therapy and Research, 3, 349-354.

Collins, B.E., Martin, J.C., Ashmore, R.D., and Ross, L. (1974). Some dimensions of the internal-external metaphor in theories of personality. Journal of Personality and Social Psychology, 29, 381-391.

Cooper, H.M., and Burger, J.M. (1980). How teachers explain students' academic performance: A categorization of free response academic attributions. American Educational Research Journal, 17, 95-109.

Cosentino, F., and Heilbrum, A.B. (1964). Anxiety correlates of sex-role identity in college students. Psychological Reports, 14, 729-730.

Covington, M.V., & Omelich, C.L. (1979). Are causal attributions causal? A path analysis of the cognitive model of achievement motivation. Journal of Personality and Social Psychology, 37, 1487-1504.

DeCharms, R. (1978). Personal causation: The internal affective

determinants of behavior. New York: Academic Press.

Elig, T.W., and Frieze, I.H. (1979). Measuring causal attributions for success and failure. Journal of Personality and Social Psychology, 37, 621-634.

Feltz, D.L. (1982). Path analysis of the causal elements in Bandura's theory of self-efficacy and and anxiety-based model of avoidance behavior. Journal of Personality and Social Psychology, 42, 764-781.

Feltz, D.L., Landers, D.M., and Raeder, U. (1979). Enhancing self-efficacy in high avoidance motor tasks: A comparison of modeling techniques. Journal of Sport Psychology, 1, 112-122.

Festinger, L. (1954). A theory of social comparison processes. Human Relations, 7, 117-140.

Friedman, M., & Rosenman, R.H. (1974). Type A behavior and your heart. New York: Knopf.

Frieze, I.H. (1976). Causal attributions and information seeking to explain success and failure. Journal of Research in Personality, 10, 293-305.

Frieze, I.H., and Snyder, H.N. (1980). Children's beliefs about the causes of success and failure in school settings. Journal of Educational Psychology, 72, 186-196.

Gall, M.D. (1969). The relationship between masculinity-femininity and manifest anxiety. Journal of Clinical Psychology, 25, 294-295.

Garber, J., & Seligman, M.E.P. (1980). Human helplessness: Theory and applications. New York: Academic Press.

Gist, M.E. (1987). Self-efficacy: Implications for organizational behavior and human resource management, Academy of Management Review, 12, 472-485.

Glass, D.C. Behavior patterns, stress, and coronary disease. Hillsdale, N.J.: Erlbaum.

Gould, J., Weiss, M. & Weinberg, R. (1981). Psychological characteristics of successful and nonsuccessful Big Ten wrestlers, Journal of Sports Psychology, 3, 69-81.

Gray, S.W. (1957). Masculinity-femininity in relation to anxiety and social acceptance. Child Development, 28, 203-214.

Grimm, L.G., and Yarnold, P.R. (1984). Performance standards and the Type A behavior pattern. Cognitive Therapy and Research, 8, 59-66.

Hartford, T.C., Willis, C.H., and Deabler, H.L. (1967). Personality correlates of masculinity-femininity. Psychological Reports, 21, 881-884.

Heider, F. (1958). The psychology of interpersonal relations. New York: Wiley.

Highlen, P.S., and Bennett, B.B. (1979). Psychological characteristics of successful and unsuccessful elite wrestlers: An exploratory study. Journal of Sport Psychology, 3, 69-81.

Hollandsworth, J.G., Jr., Glazeski, R.C., Kirkland, K., Jones, G.E., and van Norman, L.R. (1979). An analysis of the nature and effects of test anxiety: Cognitive, behavioral, and physiological components. Cognitive Therapy and Research, 3, 165-180.

Issac, S., and Michael, W.B. (1981). Handbook in research and

evaluation. San Diego, CA: Edits.

Jenkins, C.D., Zyzanski, S.J., and Rosenman, R.H. (1971). Progress toward validation of a computer-scored test for the Type A coronary-prone behavior pattern. Psychosomatic Medicine, 33, 193-201.

Jung, C.G. (1953). Anima and animus. In Two essays on clinical psychology: Collected works of C.G. Jung. (Vol. 7, pp. 186-209), Bollinger Foundation.

Kahneman, D., and Tversky, A. (1973). On the psychology of prediction. Psychological Review, 80, 237-251.

Kazdin, A.E. (1979). Imagery elaboration and self-efficacy in the covert modeling treatment of unassertive behavior. Journal of Counseling and Clinical Psychology, 47, 725-733.

Kelley, H.H. (1967). Attribution theory in social psychology. In Nebraska Symposium on Motivation (Vol. 15). Lincoln: University of Nebraska Press.

Kelley, H.H. (1971). Attributions in social interactions. Morristown, New Jersey: General Learning Press.

Krantz, D.S., Glass, D.C., & Snyder, M.L. (1974). Helplessness, stress level, and the coronary-prone behavior pattern. Journal of Experimental Social Psychology, 10, 284-300.

Lefcourt, H.M. (1976). Locus of control: Current trends in theory and research. Hillsdale, NJ: Erlbaum.

Lent, R.W., and Larkin, K.C. (1984). Relation of self-efficacy expectations to academic achievement and persistence. Journaal of Counseling Psychology, 31, 356-362.

Lent, R.W., Brown, S.D., and Larkin, K.C. (1986). Self-efficacy in the prediction of academic performance and perceived career options. Journal of Counseling Psychology, 37, 265-269.

Liebert, R.M., and Morris, L.W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. Psychological Reports, 20, 975-979.

Locke, E.A., Frederick, E. Lee, C., & Bobko, P. (1984). Effect of self-efficacy, goals, and task strategies on task performance. Journal of Applied Psychology, 69, 241-251.

Macoby, E.E. (1966). Sex differences in intellectual functioning. In E.E. Macoby (Ed.), The development of sex differences (pp. 25-55). Stanford, CA: Stanford University Press.

Mahoney, M.J., and Avener, M. (1977). Psychology of the elite athlete: An exploratory study. Cognitive Therapy and Research, 1, 135-141.

Mathews, K.A., Krantz, D.S., Dembroski, T.M., and MacDougall, J.M. (1982). The unique and common variance in the Structured Interview and the Jenkins Activity Survey measures of the Type A behavior pattern. Journal of Personality and Social Psychology, 42, 303-313.

McMahn, I.D. (1973). Relationships between causal attributions and expectancy of success. Journal of Personality and Social Psychology, 28, 108-115.

Mitchell, L.K., and Krumboltz, J.D. (1984). Research on human decision making: Implications for career decision making and counseling. In S.D. Brown & W.R. Lent (Eds.), Handbook of Counseling

Psychology (pp. 238-280). New York: Wiley.

Molnar, G.E., and DeLauretis, R.J. (1973). Predicting the curriculum mobility of engineering students: A comparison of discriminant procedures. Journal of Counseling Psychology, 20, 50-59.

National Institute of Health, Review Panel of Coronary Prone Behavior and Coronary Heart Disease. (1981). Coronary-prone behavior and coronary heart disease: A critical review. Circulation, 63, 1199-1215.

Nelson, L.R., and Furst, M.L. (1972). An objective study of the effects of expectation on cognitive performance. Journal of Psychology, 81, 69-72.

Parkerson, J.A., Lomax, R.G., Schiller, D.P., and Walberg, H.J. (1984). Exploring causal models of educational achievement. Journal of Educational Psychology, 76, 638-646.

Parsons, R. and Bales, R.F. (1955). Family, socialization and interaction process. New York: Free Press.

Paulman, R.G., and Kennelly, K.J. (1984). Test anxiety and ineffective test taking: Different names, same construct? Journal of Educational Psychology, 76, 279-288.

Naveh-Benjamin, M., McKeachie, W.J., and Yi-Guang, L. (1987). Two types of test-anxious students: Support for an information processing model. Journal of Educational Psychology, 79, 131-136.

Nielson, W.R., and Dobson, K.S. (1980). The coronary-prone behavior pattern and trait anxiety: Evidence for discriminant validity. Journal of Consulting and Clinical Psychology, 48, 546-547.

Perimuter, L.C., & Monty, R.A. (Eds.). (1979). Choice and perceived control. Hillsdale, N.J.: Erlbaum.

Porac, J.F. (1981). Causal loops and other intercausal perceptions in attributions for exam performance. Journal of Educational Psychology, 73, 587-601.

Rosenbaum, R.M. (1972). A dimensional analysis of the perceived causes of success and failure. Unpublished doctoral dissertation, University of California, Los Angeles.

Rosenman, R.H., Friedman, M., Strauss, R., Wurm, M., Kositchek, R., Hahn, W., and Werthessen, N.T. (1964). A predictive study of coronary heart disease. Journal of the American Medical Association, 189, 15-22.

Rotter, J.B., Chance, J.E., & Phares, E.J. (1972). Applications of a social learning theory of personality. New York: Holt, Rinehart, & Winston.

Schunk, D.H. (1984). Self-efficacy perspective on achievement behavior. Educational Psychologist, 19, 48-58.

Schunk, D.H. (1985). Self-efficacy and classroom learning. Psychology in the Schools, 22, 208-223.

Schunk, D.H., & Gunn, T.P. (1986). Self-efficacy and skill development: Influence of task strategies and attributions. Journal of Educational Research, 79, 238-244.

Sears, R.R. (1970). Relations of early socializationexperiences to self-concepts and gender role in middle childhood. Childhood Development, 41, 267-289.

Siegel, R.G., Galssi, J.P. & Ware, W.B. (1985). A comparison of two

models for predicting mathematics performance: Social learning versus math aptitude-anxiety. Journal of Counseling Psychology, 32, 531-538.

Spielberger, C.D. (1972). Conceptual and methodological issues in anxiety research. In C.D. Spielberger (Ed.), Anxiety: Current trends in theory and research (Vol. 2). New York: Academic Press.

Spielberger, C.D. (1980). Test Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press, Inc.

Spielberger, C.D., Gorsuch, R.L., and Lushene, R.E. Manual for the State-Trait Anxiety Inventory. Palo Alto, California: Consulting Psychologists Press, Inc.

Spielberger, C.D., Gonzalez, H.P., Taylor, C,J,, Algaze, B., and Anton, W.D. (1978). Examination stress and test anxiety. In C.D. Spielberger and I.G. Sarason (Eds.), Stress and anxiety (Vol. 5). New York: John Wiley & Sons.

Suls, J.M., and Miller, R.I. (Eds.). (1977). Social comparison processes: Theoretical and empirical perspectives. Washington, D.C.: Hemisphere.

Taylor, J. (1987). Predicting athletic performance with self-confidence and somatic and cognitive anxiety as a function of motor and physiological requirements in six sports. Journal of Personality, 55, 139-153.

Taylor, M.S., Locke, E.A., Lee, C., and Gist, M.E. (1984). Type A behavior and faculty research productivity: What are the mechanisms? Organizational Behavior and Human Performance, 34, 402-418.

Terborg, J.R. (1977). Validation and extension of an individual differences model of work performance. Organizational Behavior and Human Performance, 18, 188-216.

Triandis, H.C. (1972). The analysis of subjective culture. New York: Wiley.

Van Egeren, L.F. (1979). Cardiovascular changes during social competition in a mixed-motive game. Journal of Personality and Social Psychology, 37, 858-864.

Walberg, H.J. (1981). A theory of educational productivity. In F.H. Farley and N. Gordon (Eds.) Psychology and Education: The state of the union (pp.36 - 77). Berkeley, CA: McCutchan.

Ward, C.H., & Eisler, R.M. (1987). Type A behavior, achievement striving, and a dysfunctional self-evaluation system. Journal of Personality and Social Psychologyu, 53, 3118-326.

Webb, A.P. (1963). Sex-role preferences and adjustment in early adolescents. Child Development, 43, 609-618.

Weidner, G., & Mathews, K.A. (1978). Reported physical symptoms elicited by unpredictable events and the Type A coronary-prone behavior pattern. Journal of Personality and Social Psychology, 36, 1213-1220.

Weinberg, R.S., Gould, D., and Jackson, A. (1979). Expectations and performance: An empirical test of Bandura's self-efficacy theory. Journal of Sport Psychology, 1, 320-331.

Weinberg, R.S., Gould, D., Yukelson, D. and Jackson, A. (1981). The effect of preexisting and manipulated self-efficacy on competitive muscular endurance task. Journal of Sport

Psychology, 2, 340-349.

Weinberg, R.S., Yukelson, D., and Jackson, A. (1980). Effects of public and private efficacy expectations on competitive performance. Journal of Sport Psychology, 2, 340-349.

Weiner, B. (1974). An attributional interpretation of expectancy-value theory. In B. Weiner (Ed.), Cognitive views of human motivation (pp. 51-69). New York: Academic Press.

Weiner, B. (1979). A theory of motivation for some classroom experiences. Journal of Educational Psychology, 71, 3-25.

Weiner, B. (1980). A cognitive (attribution)-emotion-action model of motivated behavior: An analysis of judgments of help-giving. Journal of Personality and Social Psychology, 39, 186-200.

Weiner, B. (1983). Some methodological pitfalls in attributional research. Journal of Educational Psychology, 75, 530-543.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. Psychological Review, 92, 548-573.

Weiner, B., Nierenberg, R., and Goldstein, M. (1976). Social learning (locus of control) versus attributional (causal stability) interpretations of expectancy of success. Journal of Personality, 44, 52-68.

Weiner, B., Frieze, I.H., Kukla, A., Reed, L. Rest, S. and Rosenbaum, R.M. (1971). Perceiving the causes of success and failure. Morristown, NJ: General Learning Press.

Wheeler, K.G. (1983). Comparisons of self-efficacy and expectancy models of occupational preferences for college males and females. Journal of Occupational Psychology, 56, 73-78.

White, R.W. (1959). Motivation reconsidered: The concept of competence. Psychological Review, 66, 297-333.

Wilson, V.L., and Palmer, D.J. (1983). Latent partition analysis of attributions for actual achievement. American Educational Research Journal, 20, 581-589.

Wilson, V.L., and Palmer, D.J. (1983). Latent partition analysis of attributions for actual achievement. American Educational Research Journal, 20, 581-589.

Wright, L. (1988). The Type A behavior pattern and coronary heart disease. American Psychologist, 43, 2-14.

APPENDIXES

APPENDIX A

EXAM PREDICTION QUESTIONNAIRE

1. I predict that I will answer _____ (from 0 to 40) questions correctly on this test. Only one number should be selected.

2. How confident are you of this prediction (only circle one)?

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

3. In predicting your test score, how influential were the following factors listed below? Read each factor carefully and score each factor by circling only one number in the following way:

1 - Of no influence
2 - Of very slight influence
3 - Of slight influence
4 - Of medium influence
5 - Of high influence
6 - Of very high influence
7 - Of extremely high influence

A. How well the teacher presented the material for this test.	1	2	3	4	5	6	7
B. How other students who appear similar to me are doing.	1	2	3	4	5	6	7
C. How difficult I think this test will be.	1	2	3	4	5	6	7
D. My self-confidence level right now.	1	2	3	4	5	6	7
E. My concentration level while studying.	1	2	3	4	5	6	7
F. My general academic ability.	1	2	3	4	5	6	7
G. My anxiety level right now.	1	2	3	4	5	6	7
H. My knowledge of the material tested.	1	2	3	4	5	6	7
I. How well I have done in other courses.	1	2	3	4	5	6	7
J. How interesting I find the material.	1	2	3	4	5	6	7
K. My mood right now.	1	2	3	4	5	6	7
L. The amount of encouragement given to me by the teacher.	1	2	3	4	5	6	7
M. My physical health right now.	1	2	3	4	5	6	7
N. The effectiveness of my study skills.	1	2	3	4	5	6	7
O. How lucky I feel right now.	1	2	3	4	5	6	7
P. The amount of effort spent studying.	1	2	3	4	5	6	7
Q. Previous test performance in this course.	1	2	3	4	5	6	7

4. - In column A check the four most important sources of information that you used in making your test prediction.
- Once these four items have been selected, rank them from 1 (most important) to 4 (least important) in Column B.

(A) (B)

___ ___ How well the teacher presented the test material.

___ ___ How other students who appear similar to me are doing.

___ ___ How difficult I think this test will be.

___ ___ My self-confidence level right now.

___ ___ My concentration level while studying.

___ ___ My general academic ability.

___ ___ My anxiety level right now.

___ ___ My knowledge of the material tested.

___ ___ How well I have done in other courses.

___ ___ How interesting I find the material.

___ ___ My mood right now.

___ ___ Amount of encouragement given by the teacher.

___ ___ My physical health right now.

___ ___ Effectiveness of my study skills.

___ ___ How lucky I feel right now.

___ ___ Amount of effort exerted while studying.

___ ___ Performance on previous tests in this course.

APPENDIX B

TABLES

TABLE 1

DESCRIPTIVE STATISTICS

Item	TEST 1 Range	Mean	Std. Dev.
Predicted Test Score	20 to 40	31.68	3.5
Actual Test Score	21 to 40	31.27	4.17
Accuracy Score	-12 to 11	-.45	3.76
Confidence of Prediction	30 to 100%	70.83%	16.31%

Item	TEST 2 Range	Mean	Std. Dev.
Predicted Test Score	15 to 40	31.32	3.5
Actual Test Score	19 to 40	30.76	4.17
Accuracy Score	-12 to 8	-.59	3.76
Confidence of Prediction	20 to 100%	67.52%	16.33%

TABLE 2 (Continued)

Item	TEST 3 Range	Mean	Std. Dev.
Predicted Test Score	20 to 40	32.01	3.58
Actual Test Score	14 to 39	30.94	4.09
Accuracy Score	-18 to 9	-.95	4
Confidence of Prediction	30 to 100%	69.36%	16.26%

Item	TEST 4 Range	Mean	Std. Dev.
Predicted Test Score	25 to 40	32.5	3.18
Actual Test Score	19 to 39	27.97	4.14
Accuracy Score	-14 to 6	-2.6	3.78
Confidence of Prediction	20 to 100%	70.38%	16.05%

TABLE 2

PEARSON CORRELATION COEFFICIENTS AMONG PREDICTED TEST SCORE (PTS), ACTUAL TEST SCORE (ATS), CONFIDENCE OF PREDICTION (CP), AND ACCURACY SCORE (AS)

CORRELATION	Test 1	Test 2	Test 3	Test 4
PTS/ATS	.531**	.464**	.522**	.523**
CP/AS	.016	.028	.014	.061
CP/PTS	.362**	.178*	.285**	.263*

* P < .05, ** P < .01

TABLE 3

MEAN RATINGS OF EFFICACY INFORMATION SOURCES
PART 3 OF EXAM PREDICTION QUESTIONNAIRE

Information Source	Test 1	Test 2	Test 3	Test 4
Effort	5.197	5.102	5.178	5.102
Knowledge	5.491	5.459	5.401	5.561
Test Difficulty	5.101	4.981	4.643	4.535
Study Skills	4.924	4.962	4.873	5.001
Teacher Effectiveness	4.586	4.573	4.217	4.242
Performance of Others	2.427	2.465	2.369	2.471
Concentration Level	5.141	5.229	5.153	5.025
Academic Ability	4.822	4.491	4.492	4.631
Test Performance	No Data	4.287	4.395	4.643
Other Class Grades	4.408	3.561	3.376	3.669
Interest Level	4.389	4.242	4.229	4.306
Teacher Encouragement	2.987	3.025	2.975	3.013
Self-confidence Level	5.153	4.924	4.701	4.931
Anxiety Level	4.452	4.427	3.955	4.115
Mood	4.433	4.401	3.968	3.975
Physical Health	3.968	3.682	3.071	3.605
Luck	2.904	3.146	3.071	3.025

TABLE 4

MEAN RATINGS OF EFFICACY CUES BY GENERAL GROUP
PART 3 OF EXAM PREDICTION QUESTIONNAIRE

Group Criteria	Test 1	Test 2	Test 3	Test 4
Task-preparation	4.86	4.93	4.92	4.87
Performance-related	4.19	3.97	3.85	3.99
Personal feelings	3.92	3.93	3.72	3.78

TABLE 5

CUMULATIVE RATINGS OF EFFICACY CUES FROM
PART 4 OF EXAM PREDICTION QUESTIONNAIRE

Information Source	Test 1	Test 2	Test 3	Test 4
Effort	255	284	277	279
Knowledge	336	278	326	378
Test Difficulty	154	130	102	79
Study Skills	124	136	104	119
Teacher Effectiveness	119	100	87	63
Performance of Others	7	9	0	0
Concentration Level	176	197	205	170
Academic Ability	93	45	55	56
Test Performance	No Data	65	93	76
Other Class Grades	53	24	21	21
Interest Level	67	53	52	39
Teacher Encouragement	0	0	0	0
Self-confidence Level	58	76	81	72
Anxiety Level	50	52	57	43
Mood	16	27	26	23
Physical Health	33	26	10	23
Luck	14	15	8	12

TABLE 6

RANKINGS OF EFFICACY CUES FROM 1 (HIGH) TO 17 (LOW)
PART 4 OF EXAM PREDICTION QUESTIONNAIRE

Information Source	Test 1	Test 2	Test 3	Test 4	Mean
Effort	2	1	2	2	2
Knowledge	1	2	1	1	1
Test Difficulty	4	5	5	5	5
Study Skills	5	4	4	4	4
Teacher Effectiveness	6	6	7	8	7
Performance of Others	15	16	16	16	16
Concentration Level	3	3	3	3	3
Academic Ability	7	11	10	9	9
Test Performance	No Data	8	6	6	6
Other Class Grades	10	14	13	14	13
Interest Level	8	9	11	11	10
Teacher Encouragement	16	17	17	17	17
Self-confidence Level	9	7	8	7	8
Anxiety Level	11	10	9	10	11
Mood	13	12	12	12	12
Physical Health	12	13	14	13	14
Luck	14	15	15	15	15

TABLE 7

MEAN RATINGS OF EFFICACY CUES BY GENERAL GROUP
PART 4 OF EXAM PREDICTION QUESTIONNAIRE

Group Criteria	Test 1	Test 2	Test 3	Test 4
Task-preparation	179.5	174.7	175.2	174.7
Performance-related	76.7	54.6	54.2	46.4
Personal feelings	28.5	20.1	29.5	28.8

TABLE 8

POINT BISERIAL CORRELATION COEFFICIENTS
PART 4 OF EXAM PREDICTION QUESTIONNAIRE

Information Source	Test 1	Test 2	Test 3	Test 4
Effort	-.151	-.037	-.017	.112
Knowledge	.044	-.041	.103	.089
Test Difficulty	.029	.183	-.139	-.091
Study Skills	.129	-.113	.002	.051
Teacher Effectiveness	-.055	.023	.004	.018
Performance of Others	-.147	.015	-.112	.013
Concentration Level	-.003	.104	.015	.074
Academic Ability	.098	-.041	-.006	-.019
Test Performance	No Data	-.122	-.018	-.101
Other Class Grades	-.094	.071	.026	.097
Interest Level	-.015	-.094	.056	-.043
Teacher Encouragement	.021	-.046	.033	-.029
Self-confidence Level	.061	.096	.011	-.125
Anxiety Level	.072	-.019	.025	-.137
Mood	-.063	-.121	-.185	-.129
Physical Health	-.177	.067	.071	-.021
Luck	.004	.026	.093	.018

* P < .05, * P < .01

TABLE 9

MULTIPLE REGRESSION EQUATION

Multiple R: .756 Multiple R^2: .572 Adjusted Multiple R^2: .548

INDEPENDENT VARIABLE	COEFFICIENT	T	P (2 Tail)
Constant	-45.520	-5.272	0.000
Age	-.127	-1.004	0.317
Gender	-1.681	-1.061	0.291
Year	-0.614	-0.586	0.559
GPA	3.526	2.247	0.026
JAS	-1.80	-6.717	0.001
TAI	0.099	1.826	0.070
BSRI	0.054	0.814	0.417
Test Scores	0.321	6.157	0.001

TABLE 9 (Continued)

Analysis of Variance

Source	Sum-of-Squares	DF	Mean-Square	F-Ratio	P
Regression	10605.79	8	1325.72	24.68	0.001
Residual	7949.93	148	53.72		

APPENDIX C

FIGURES

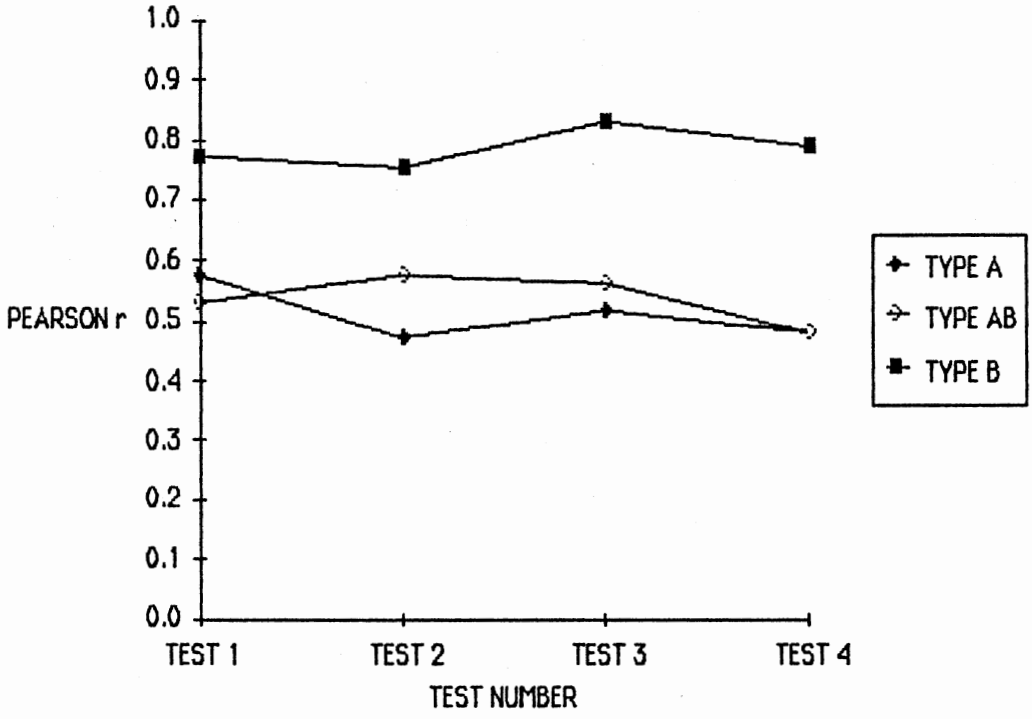


Figure 1. Predicted test score/actual test score correlations as a function of the Type A - Type B behavior pattern.

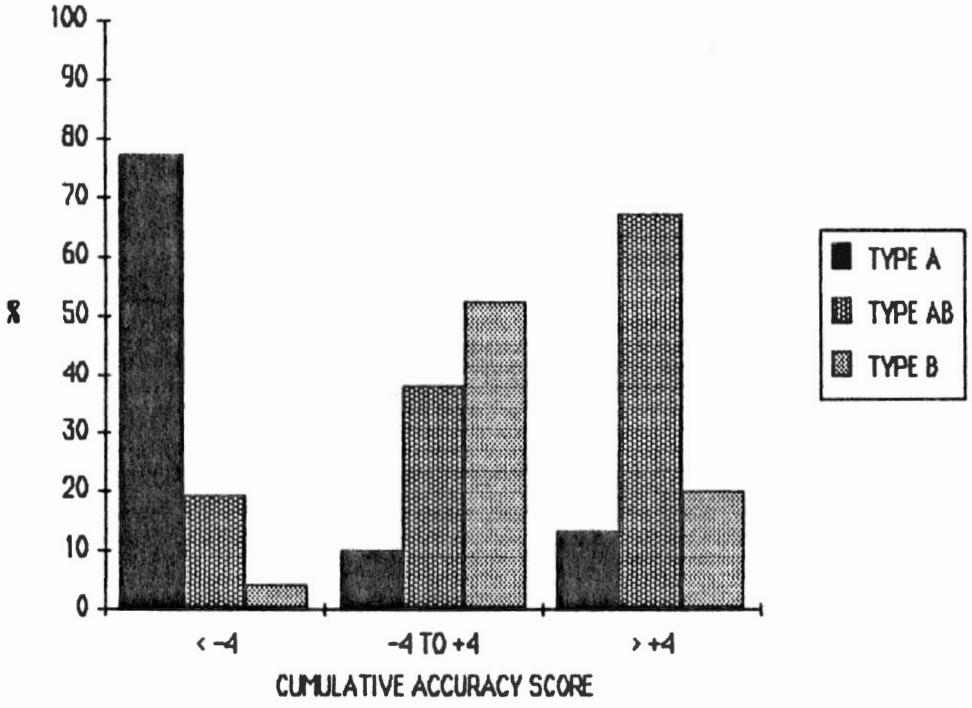


Figure 2. Cumulative accuracy scores as a function of the Type A - Type B behavior pattern.

VITA

Gerald L. Harbour

Candidate for the Degree of

Doctor of Philosophy

Thesis: FACTORS AFFECTING THE ACCURACY OF EFFICACY JUDGMENTS

Major Field: Applied Behavioral Studies

Biographical:

Personal Data: Born in Everett, Washington, March 31, 1949, the son of Joe and Ione Harbour.

Education: Graduated from Cascade High School, Everett, Washington, in June, 1967; received Bachelor of Arts Degree in Geology from Western Washington State College in December, 1972; received Master of Science Degree in Geology from Eastern Washington State College in June, 1978; completed requirements for the Doctor of Philosophy degree at Oklahoma State University in December, 1988.

Professional Experience: Geologist, Gulf Resource and Chemical Corporation, March, 1975 to May, 1978; Geologist, Idaho Bureau of Mines and Geology, June 1978 to May, 1980; Geological Supervisor, Phillips Petroleum Company, June, 1980 to May, 1986; Educational Researcher, Child Service Demonstration Center, June, 1988 to December, 1988.