Old Dominion University
ODU Digital Commons

Psychology Theses & Dissertations
Psychology

Fall 1991

Multiple Cue Probability Learning: Effects of Individual Differences on MCPL Performance

Catherine Q. Greenwald
Old Dominion University

Follow this and additional works at: https://digitalcommons.odu.edu/psychology_etds

Part of the Psychology Commons

Recommended Citation

Greenwald, Catherine Q.. "Multiple Cue Probability Learning: Effects of Individual Differences on MCPL Performance" (1991). Master of Science (MS), Thesis, Psychology, Old Dominion University, DOI: 10.25777/86rn-nf42
https://digitalcommons.odu.edu/psychology_etds/596

This Thesis is brought to you for free and open access by the Psychology at ODU Digital Commons. It has been accepted for inclusion in Psychology Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

MULTIPLE CUE PROBABILITY LEARNING: EFFECTS OF INDIVIDUAL DIFFERENCES ON MCPL PERFORMANCE

by

Catherine Q. Greenwald
B.A. May 1989, The Pennsylvania State University

A Thesis Submitted to the Faculty of Old Dominion University in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

PSYCHOLOGY

OLD DOMINION UNIVERSITY
December, 1991

Approved by:

Robert M. McIntyre, ODU
Chair

Mark M. Scerbo, ODU

Michelle L. Kelley, ODU

ABSTRACT

Multiple Cue Probability Learning: Effects of Individual Differences on MCPL Performance

Catherine Q. Greenwald
Old Dominion University, 1991
Chair: Dr. Robert M. McIntyre

This exploratory study investigated individual difference effects on multiple cue probability learning performance. Measures of cognitive complexity, creativity, problem solving ability, and anxiety were correlated with the MCPL indices of performance. Eighty participants (40 men and 40 women) were randomly selected graduate students from the six colleges of Old Dominion University. Results showed significant Pearson Product Moment correlations among the Level 2 MCPL achievement and matching indices and the predictors Career Path Appreciation phrase card average, Kirton Adaptation Innovation Inventory, Quantitative GRE scores, gender, and anxiety. There is some evidence that MCPL performance can be explained in terms of these variables, but that evidence is not definitive because of study limitations. This evidence suggests the beginning stages of a theory of MCPL performance. Suggestions for future research include counterbalancing the order of MCPL Levels 2 and 3 in order to counteract fatigue effects, improving Mastermind scoring procedures, and continuing to investigate effects of GRE scores.

ACKNOWLEDGMENTS

This thesis was accomplished with the assistance of several people who deserve my gratitude in print. The members of my committee provided support and insights that made this document the best it could be. I would like to thank Dr. Michelle Kelley for providing immediate and supportive feedback and suggestions during the many stages of this thesis. I am grateful to Dr. Mark Scerbo for his critical analysis of my design and his quick return of my drafts. Dr. Robert McIntyre provided more time and effort for my thesis than I ever expected. I thank him for the endless hours he spent helping me get it right and for his insistence on perfection. There are many moments I could never adequately thank him for.

I am also grateful to my reader, Dr. Raymond Kirby, for the immense help he provided when I needed it at the last minute. I would like to thank Dr. T. O. Jacobs for the hundreds of hours of videotape scoring he provided under less than perfect conditions. The data collection could not have been completed without his remarkable effort. Thanks also to Pamela Jordan and Laura Hamill, who gave their time to assist with the participant interviews. I would also like to thank Chris Rotolo, who helped me during the beginning stages of this project.

A final thanks to the fellow graduate students who gave up several hours of their time to help me finish my thesis

by participating.

This thesis is dedicated to my parents who have encouraged me through the years and provided me with everything I needed to get this far. I would also like to dedicate it to Bill Mergen and Lora Jacobi who supported and encouraged me in my darkest hours and rejoiced in my success.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Multiple Cue Probability Learning: Effects of Individual Differences on MCPL Performance

I. INTRODUCTION

Research on multiple cue probability learning (MCPL) has been conducted for over twenty years. MCPL research has concentrated on manipulating time pressure, task complexity, feedback, or training to investigate effects on an individual's judgments about linear and curvilinear relationships between two or more probabilistic cues (Brehmer, 1972; Hammond, Summers & Deane, 1973; Rothstein, 1986). But despite the number of studies conducted in the area, individual differences in processes used to make the inferences, or abilities correlated with the skill to form and test hypotheses about probabalistic cues, have not been investigated. The purpose of the present research is to add to the current understanding of individual differences in performance on the MCPL task and to initiate a construct validation of performance on the MCPL task. The question under investigation can be summarized as follows: Are there patterns of individual differences in the performance of MCPL tasks?

Multiple-cue probability learning task

The multiple cue probability learning task requires that an individual make a decision or judgment about some recurring event on the basis of cues presented to the person. Over time, and through trial and error, the decision maker may learn to use the multiple cues appropriately or at least consistently to predict the actual event correctly. Much of the research on MCPL takes place in the laboratory with artificial tasks. However, in real life, there are many examples of multiple cue probability learning tasks. For example, consider weather prediction. Based on several cues (barometric pressure, temperature gradients, wind shifts, and so forth) the weather forecaster learns over time (or, alternatively, over a series of trials) to make judgments about upcoming weather patterns and the probability of a particular outcome (such as a thunderstorm). In the field of mental health, a clinician learns over time to diagnose specific disorders after observing many people who display a set of symptoms.

In the research laboratory, multiple-cue probability learning tasks can be presented to research participants by means of a computer. Rothstein's task (1986) is an example of this approach. A participant is presented with the task on a video monitor which shows two vertical bars as cues, whose height varies on a scale from one to ten. The participant's job is to judge the height of a third bar based on the cues. Once the participant forms a judgment, he or she adjusts a third column to the estimated height (again one to ten). After this judgment is made, a vertical "feedback" bar of the correct height is presented to the participant.

Slovic, Fischhoff, and Lichtenstein (1977) presented a review of the behavioral decision theory literature which describes ways individuals use multiple cues to predict events. This review indicates that, in general, people search for and test hypotheses about the cue-criterion relations. The review also indicates that a person learns to perform successfully on the task more readily when there are linear relationships rather than nonlinear relationships among the cues and between the cues and criterion. When nonlinear relations are introduced, people learn more slowly, test hypotheses more inconsistently, and use outcome feedback inefficiently. Participants may forget which rules they have tested and consequently resample rules they have already discarded instead of creating unique hypotheses (Slovic, Fischhoff, & Lichtenstein, 1977).

One possible reason for the difference in a person's performance on linear and nonlinear relationships is that linear relations are the most commonly encountered from earliest years in grade school to adulthood. Nonlinear relationships are not necessarily learned in grade school, if ever, and are less accessible. This idea, if true, has interesting implications regarding hypothesized relationships between various experiences and MCPL performance. These will be presented below.

The typical responses to the MCPL task described by Slovic and colleagues have led Klayman (1988) to suggest the skills necessary for mastery of MCPL tasks. Klayman states that research participants must learn the function relating each cue to the criterion, the relative weights of the cues, and the best way to combine the information from several cues into a single judgment. A method that has been suggested as a model for describing the way people do this is Brunswick's lens model.

Brunswick's Lens Model

Wiggins (1973) refers to Brunswick's "probabilistic functionalism" as relevant to the way clinicians make diagnostic decisions. Brunswick's lens model provides a means of explaining how decision makers use probabalistic cues (that is, cues that have a probabalistic rather than a deterministic relationship to the outcome of interest) to make judgments about that outcome. Figure 1 illustrates the lens model as applied to the present study. The cues are symbolized along the center of the lens. In a sense, these form the lens through which the decision maker views and decides on the environmental event on the right side of the lens. On the right side is the judge's prediction based on the cues. On the left, is the actual environmental event about which decisions are made. The model is useful because it provides a framework for describing how the decision maker uses cues to make a judgment. The model also suggests statistical procedures for representing a single judge's decision processes. Multiple regression procedures are used to describe a paramorphic, linear representation of the decision maker's policy.

In their discussion of Brunswick's lens model, Dudycha and Naylor (1966) described several statistical indices.

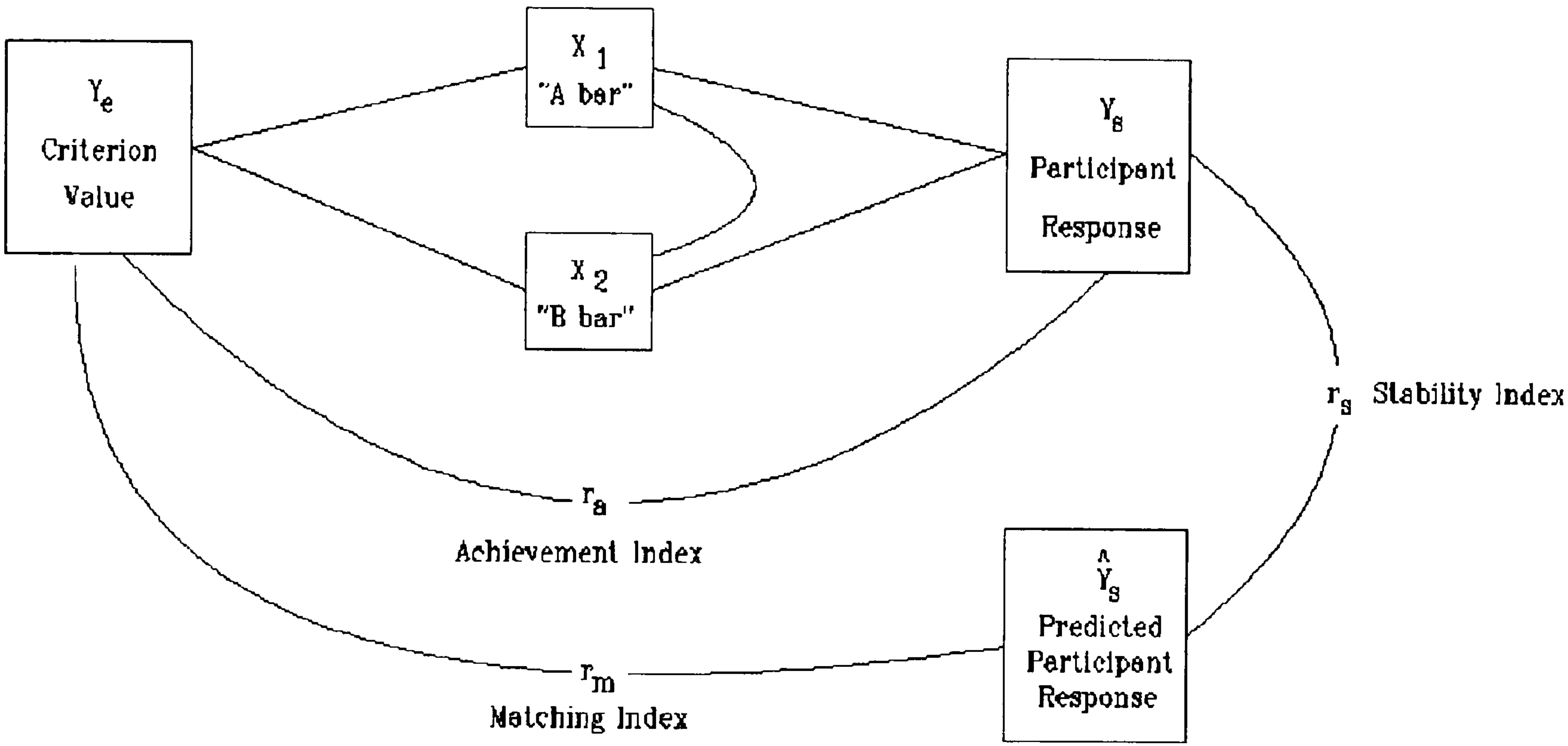


Figure 1. A representation of the Brunswick Lens Model for a multiple cue probability learning task.

The first, the index of environmental predictability (r_e), is the correlation between the predicted environmental event (Y'_e) and the actual event (Y_e). It is equivalent to R_e, the multiple correlation between Y_e and the set of cues. The important indices representing the participant's performance for the present research are r_a, r_m, and r_s.

Second, Dudycha and Naylor describe the achievement index (r_a) as the correlation between the judge's decision (Y_s) and the actual environmental event being predicted (Y_e). Thus, this index measures the degree of agreement between the criterion values and responses of the participant over 100 trials.

The third is the matching coefficient (r_m), which indicates the degree to which the participant's strategy or policy matches the optimal criterion equation. If a participants' matching policy is identical to the optimal policy equation, r_m is equal to 1.0. The matching coefficient (r_m) is defined as the correlation between the predicted value of the decision maker's response (Y'_s) and the predicted value of the actual event (Y'_e). Thus the matching coefficient is essentially the achievement index when r_a is corrected for the unreliability of the participants' policy (Y'_s).

The fourth index is that of subject consistency (r_s, or equivalently, R_s). The index of subject consistency measures the degree to which the participant consistently uses his or her strategy across the trials. It is the multiple correlation in which the decision maker's vector of judgments over many trials is the criterion and the cues are the predictors. In other words, r_s is the correlation between the vector of decisions that the decision maker makes (Y_s) and the optimal linear composite of these cues (Y'_s).

The focus of the present research is to identify individual differences that explain variability in performance on a MCPL task when performance is defined in terms of the three indices defined above. How participants make and test hypotheses about the cue-criterion relations, and the individual differences that may affect those processes is not addressed conclusively in the MCPL literature, but there are clues pointing to identifiable individual differences.

Individual Differences

Brehmer's research has made a major contribution to the present knowledge about multiple cue probability learning and the external factors that affect performance. His recent work (1985) reported that linear relations are learned faster than nonlinear relations. Brehmer suggested that perhaps not all people have the capacity to use nonlinear hypotheses appropriately because results showed that even people who had relevant inverted-U rules in a pre-task hypothesis assessment could not always learn the nonlinear (inverted-U rule) task. In his research, Brehmer asked participants to provide as many hypotheses about relations between two scaled variables (Brehmer, Alm, & Warg, 1985) as they could before they actually tried the MCPL task. Most provided up to eight different hypotheses about the two variables (one cue and one criterion) and inverted-U rules were often mentioned.

Participants demonstrated that even if they could verbally hypothesize a nonlinear rule, they did not necessarily use that rule during the actual task. In fact, the group of participants in the Brehmer et al. study who could verbalize the nonlinear rule did not perform significantly better on the task than the group who did not verbalize the nonlinear rule. Therefore, it appears there is no transfer from verbalizing the nonlinear rules that were learned to actually applying these rules.

In 1987, Brehmer reported that although participants do well in tasks involving linear cue-criterion relations, the introduction of nonlinear cue-criterion relations seemed to produce an obstacle to learning the task. Brehmer suggested that research participants learn MCPL tasks by limiting the tested hypotheses to rules for combining cue information instead of expanding their hypotheses pool to include hypotheses based on separate information from each cue. This way of processing information imposes linear thinking, and leads research participants to test linear relations between the set of cues and criterion. Thus, learning nonlinear relations becomes a task that is rarely accomplished.

This capacity (or lack thereof) for formulating nonlinear hypotheses could be attributed to a person's construing ability. Construing ability is defined in the problem solving literature as the ability of a person to attribute probabilistic properties to events and to form ideas about the relationships between those events (Nystedt and Magnusson, 1982). Construing ability was first described by Kelly in 1955 (Landfield & Leitner, 1980; Nystedt and Magnusson, 1982) in the context of personal construct theory.

Kelly's (1980) personal construct theory describes how people construe. People anticipate events according to their personal theories that have been formed because of previous experience with similar events. It is each person's attempt to structure the onrush of occurrences by hypothesizing, experimenting, observing, and revising his or her predictions based on the combination of theory and actual observations. Kelly points out that although his assumptions are presented in a paper based on the goals of psychotherapy, "life is essentially an anticipation of events to come" (Kelly, 1980, p. 29). Thus these assumptions may be applied to other situations, such as MCPL.

The construction corollary is an elaboration of Kelly's personal construct theory and his idea that people are not shaped by events, but instead by the meaning they give to those events (Kelly, 1980). This corollary includes construing as one of four processes in which people must engage in order to develop a model of external reality.

Nystedt and Magnusson 1982) describe Kelly's construction corollary by positing these four processes: (a) people tend to anticipate events; (b) they construe properties of an ongoing stream of events; (c) they review the event as "constructed" by their perceptions, and (d) they cognitively replicate events prior to predicting. In other words, people form a broad picture of a particular type of situation based on properties that they perceive of each related event. When a similar event occurs, the person can base his or her performance (or reaction to that event) on previous experience.

The four-process corollary can be used to explain multiple cue probability learning. The external reality can be viewed as the actual relationship between multiple cues and the criterion. The cues and outcome of each trial can be considered as the events. Each person's MCPL judgments may involve the process of anticipation, construing, reviewing, and replicating. This description of MCPL performance may be more useful than the lens model because it assumes that the process of solving the task is continuous and developing, and built on previous trials instead of a trial-by-trial judgment with no effect from previous trials that is assumed by the lens model.

As mentioned above, Nystedt and Magnusson (1982) define construing ability as the ability of a person to impose meaning on probabilistic events by attributing properties to these events. Thus, a "high-construing" person might view the MCPL task from many different perspectives. For example, he or she might examine each cue, and use information from his or her observations about that cue-criterion relationship to attribute properties or relationships to future events. As was discussed before, the difficulty with nonlinear MCPL tasks is the inability of the typical decision maker to treat each of the cues separately when attributing properties to the cue-criterion relationship. Instead, the typical decision maker uses a "combination rule" to predict the environmental event which necessarily imposes linear constraints on the nonlinear task. In contrast, a person with a high construing ability may more successfully identify the related and unrelated information provided in the task. If a person has a high construing ability, his or her success at nonlinear MCPL tasks may be greater than a person with a low construing ability.

Measures of Construing Ability

One method designed to measure construing ability is the Career Path Appreciation (CPA) interview (Stamp, 1988; Jacobs & Jaques, 1989), which is described later. The CPA interview was developed by Gillian Stamp to determine the current level of capability of a person to respond to complexity in a job.

The CPA interview consists of three subtasks: A phrase card task, a symbol card sort task, and an interview about past and present work experiences. These tasks were developed by Gillian Stamp and examined in five validation studies described by Stamp (1988). A detailed explanation of the phrase card and card sort tasks, the key components of the interview, are presented in the Method section.

Stamp's concept of current level of capability is defined as "the person's capability at a particular age to generate and respond to complexity within the world, within him/herself and in the constant work needed to keep both in equilibrium" (Stamp, 1988, p. 11). The respondent reviews with the administrator the pattern of his or her working life, and reveals through the course of the interview a level of decision-making skill. This skill is similar to what has been called "construing ability." Essentially, Stamp is positing that level of construing ability is one determinant of a person's success as manager in an organization. This position has its origin in Stratified Systems Theory (SST) of organizations developed by Jacques (Stamp, 1988).

Stratified systems theory (SST) is a theory of organizations and the work within them. Associated with different positions in an organization are cognitive skill levels necessary to carry out the work of each stratum. These cognitive skills have been defined as cognitive complexity and construing ability (Jacobs & Jaques, 1989). Both phrases refer to the manner in which a worker "constructs reality" to make decisions.

SST assumes that adult workers can progress through a series of eight hierarchical levels of cognitive functioning, and that there is a predictable growth over time of capability to perform at more advanced levels. In a sense, then, SST can be thought of as representing a hierarchy of adult cognitive functioning that parallels the levels in an organization. The theory provides descriptions of the characteristics of the worker for each hierarchical level. These levels are described in Table 1.

There are three domains and seven subsystems within those domains in the hierarchy constructed by Jaques and Jacobs (1989). The highest domain, referred to as "systems", contains two subsystems. The "corporation level", the highest subsystem, involves the organization's interaction with the world environment. People capable of operating at this level of functioning are able to make long range plans for a corporation in order to guide the organization through the next 20 years. Typically people at this level are capable of developing, establishing, and dismantling business units and maintaining the organization's values.

The next subsystem level in the "systems domain" is the "group level". Workers competent at this level do well at interpreting the strategy and objectives developed by the levels above them. These workers select options for the company and determine priorities after making judgments about the world environment. The scope of their decision making is to look ahead 10 years.

The "organizational domain" represents the next step down in the hierarchy. As Table 1 indicates, there are two levels comprising this domain: "company level" and "division level". "Company level" and "division level" are concerned

Table 1

FUNCTIONAL DOMAINS IN THE REQUISITE ORGANIZATION

Time Span	Stratum	Functional Domain
20 years 10 years	VII Corporation ----------- VI Group -----------	Systems Domain -- Operates in a nearly unbounded world environment, identifies feasible futures, develops consensus on specific futures to create, and builds whole systems which can function in the environment. Conditions environment to be "friendly" to systems thus created. Creates a corporate culture and value system compatible with societal values and culture, to serve as a basis for organizational policies/climate.
5 years 2 years	V Company ----------- IV Division -----------	Organizational Domain -- Individuals at Stratum V operate bounded open systems thus created, assisted by individuals at Stratum IV in managing adaptation of those systems within the environment by modification/maintenance/fine tuning of internal processes, and climate, and by oversight of subsystems.
1 year 3 months	III Department ----------- II Section ----------- I Shop Floor	Production Domain -- Runs face-to-face (mutual recognition or mutual knowledge) subsystems -- units or groups engaged in specific differentiated functions but interdependent with other units or groups, limited by context and boundaries set within the larger system.

with long term goals not more than 5 or 2 years away, respectively. Workers competent at the "company level" are able to manage business units within the corporate policies, and contribute to the development of corporate strategies. These workers make forecasts and establish information networks. They are able to develop plans and performance appraisal systems.

Workers competent at the "division level" in the "organizational domain" are general managers who coordinate activities and develop new processes for units in order to maximize efficiency. These workers are able to allocate resources, create a supportive environment, and reinforce the motivational climate.

There are three levels of functioning within the "production domain": The "department level" workers, "section level," and "shop floor level." Workers competent at these levels have relatively short term goals to meet in accordance with directions sent down from the upper levels. Workers competent at the "department level" operate units in an effort to modify and fine tune the system in order to cope with changing trends. They are capable of establishing short-term production goals over the next two years, and work closer to the actual product and customer. Workers competent at the "section level" fall between "department level" workers and "shop floor" workers in the "production domain." These workers are the first line managers in charge of the group of workers producing the output. They are capable of looking ahead one year to plan inventory and production schedules, they provide daily production information, and develop quality control measures. Workers competent at the "shop floor" level represent the lowest level of functioning in the hierarchy, and are responsible for direct operating tasks. These workers are capable of solving immediate line problems, and look ahead no more than three months in the scope of their position.

In accord with stratified systems theory, the CPA indicates at which level on the hierarchy a worker is currently functioning. In addition, Jaques proposes a growth curve chart that predicts the level to which a worker will rise during his or her career based on current construing ability. The combined feedback of current operating level and potential operating level can help a worker and an organization to make appropriate career decisions. The hierarchical chart of cognitive complexity is a useful taxonomy because it provides guidance for matching the level of worker with the type and level of a particular job.

By means of the CPA, it is possible to categorize a person's current cognitive skill level and to predict potential for future cognitive skill level. For example, the CPA can be used to make recommendations to postpone promoting a worker who is currently functioning at one level until he or she has reached a higher level of capability based on the proposition that the worker requires more cognitive development before taking on certain responsibilities. The CPA can be used by a person to make career decisions and set goals that are realistic, while reviewing the accomplishments of the past in a suitable framework.

Stamp (1988) has hypothesized that the handling of complexity on the job, and specifically the "discretion" required, is highly related to advancement. Her longitudinal validation studies show high correlations of job complexity demands to worker capabilities as measured by the CPA. Stamp reports (1988) validity data for 182 respondents in four different studies who were followed for periods of four to thirteen years after the initial CPA interview. The criterion in this research was the level in the organization to which the individual was promoted. The predictor was the level of cognitive functioning as assessed by the CPA. Correlations range from .70 to .92 and cover a wide variety of companies (multinational oil company to a mining company in a developing nation), and a wide variety of capability levels and education levels (6th grade education to PhD level).

The Role of Creativity

In addition to the concept of cognitive complexity, the role of creativity in MCPL tasks should be considered. The concept of creative style is one that Kirton (1987) has developed over the past three decades in his adaptation-innovation theory. He argues that people vary on a continuum from adaptation to innovation and that these two styles are unrelated to cognitive capacity or creativity level. Instead, a person's creativity style, measured by the Kirton Adaptation-Innovation Inventory (KAI), indicates the degree to which a person is willing to solve problems by radical, innovative solutions on the one extreme, or by adapting to the current restrictions on the other. Neither problem solving style is considered better than the other, and neither style is associated with the level of intelligence of a person. Kirton suggests that extreme innovators and adaptors tend to engage in conflict instead of collaboration because they are unaware of their differences in problem solving style. In fact, they often seem to devalue the effectiveness of the others' style.

In an attempt to gather construct validity evidence, the KAI Manual summarizes the relationships among measures of IQ and "creativity style." Kirton (1987, p. 90) is critical of those who suggest that a minimum IQ is necessary for a person to be considered creative. He asserts that creative style (as measured by his adaptation-innovation scale) and creative level (as measured by IQ or other tests of cognitive capacity) are orthogonal to each other. Creative level is considered to be significantly related to cognitive capacity, so that a more intelligent person has a higher creative level and thus a higher quantity or quality of creative ideas. Creative style, however, is a separate dimension of a person. Creative style refers to the way a person is creative.

If cognitive style and cognitive level are orthogonal to each other as Kirton assumes, then the KAI and measures of IQ should not correlate. As reported in the Manual

(Kirton, 1987, p.91) none of the standard measures of intelligence or cognitive capacity were significantly related to scores on the KAI measure.

In addition, Kirton's factor analysis of data provided by Torrance from a 1980 study that used a variety of measures of creativity shows an emergence of two factors. Fourteen of the 15 measures correlated with either Factor I (style) or Factor II (level).

Hypotheses

Based on previous literature, and in the interests of initiating a construct validation of multiple cue probability learning task performance, hypotheses for a variety of variables are presented here.

It is hypothesized that a correlate of the MCPL task is construing ability, derived from Brehmer's (1985, 1987) suggestion that people incorrectly aggregate information from multiple cues into a single judgment. This hypothesis is that decision makers often fail to "construe" the nature of the cue-criterion relationships, but instead base their decisions on simple linear composites of the cues. Thus, there should be a positive correlation between scores on the Career Path Appreciation interview, which measures construing ability and cognitive complexity of problem solving strategies, and the different performance indices associated with performance on the MCPL task.

The Kirton Adaptation-Innovation Inventory measures problem solving styles of individuals. The two personality styles are "adaptor" and "innovator". It is not clear which problem solving style will be most effective for an MCPL task. It seems reasonable to expect that an individual identified as an "innovator" will have more flexibility when formulating hypotheses for each trial, thus performing more effectively on the task. Thus, a positive correlation among the indices of the MCPL task and the KAI score is hypothesized.

Concept-learning problem solving is another hypothesized correlate. This construct is described by Laughlin, Lange, and Adamopoulos (1982) in their research with "Mastermind." Mastermind is a concept-learning problem-solving game where the player uses problem solving strategies and unique feedback rules to identify unknown attributes of the known conceptual rule. The computer logic game used in this study is based on the commercial board game "Mastermind," and it is described in detail in the Method section. Mastermind appears to be similar to MCPL for the following reasons: 1) it involves logical thinking to discover a pattern; 2) it requires the participant to examine multiple (four) cues and make a prediction; 3) it presents trial by trial feedback on the basis that the player makes subsequent judgments to solve the puzzle. It is hypothesized that there will be a positive correlation among concept-learning problem solving performance on "Mastermind" and the indices of the MCPL task.

Demographic information regarding experience level with mathematical problem solving is hypothesized to correlate positively with MCPL performance. In addition, there are two exploratory hypotheses pertinent to demographic data. The effects of gender have not been reported in the current MCPL literature, although Brehmer (1985) suggests its use as an exploratory variable. There is some evidence in the spatial ability literature (Hyde, 1990; Meehan & Overton, 1986) that men are moderately better at spatial tasks than women. As MCPL is a spatial task, it seems appropriate to consider the possibility that there may be a positive correlation between gender and MCPL indices.

Similarly, age effects have not been investigated, but are suggested by Jacobs and Jaques (1989). Age is positively correlated with performance on the CPA interview. Therefore, if the hypothesis that participants with higher construing ability, as measured by the CPA interview, will perform better on MCPL tasks is supported, a positive correlation among age and MCPL performance is expected.

The Graduate Record Exam (GRE) measures verbal, quantitative, and analytic abilities of students applying to graduate programs. These scores show achievement in these areas, and it can be argued that higher GRE scores indicate higher general intelligence. It is hypothesized that the Verbal GRE score will positively correlate with MCPL indices because the Verbal GRE score is assumed to be a reasonable indicator of general intelligence. It is hypothesized that the Quantitative GRE score will positively correlate with the MCPL indices because a high Quantitative GRE score indicates more experience and achievement in mathematics. It is hypothesized that the Analytical GRE score will positively correlate with MCPL performance because the Analytical GRE's measure logic and puzzle solving ability, and the problems to be solved in the MCPL may be better solved by a person who performs well on puzzle solving problems.

Anxiety is another hypothesized correlate with MCPL which will be measured by the Test Anxiety Profile (TAP). The TAP is a broad ranging profile measuring two related forms of anxiety: Feelings of Anxiety (FA) and Thought Interference (TI). This measure is included in an effort to identify the extent to which participants are anxious about using a computer and participating in a psychological study. Increased anxiety about performance is expected to be negatively related to MCPL performance.

Table 2 summarizes the hypothesized covariates and their relationship to MCPL performance.

Table 2

Hypothesized Covariates			Relationship
Career Path Appreciation Interview......			positive
Kirton Adaptation-Innovation Inventory.. (higher score=innovator)			positive
Mastermind.............................			positive
Test Anxiety Profile...................			negative
Graduate Record Examinations		V.......	positive
		Q.......	positive
		A.......	positive
Demographic	Age....................		positive
	Gender.................		positive
	Number of Courses:		
	Quantitative.....		positive
	Artistic.........		positive

II. METHOD

Subjects

Research participants were 87 graduate students from selected majors at Old Dominion University in Norfolk, Virginia. Seven of the 87 were used in a pilot study, and of the other 80, 40 men and 40 women between the ages of 21 and 59 (M = 33.2, SD = 8.5) participated. All six colleges of the university were sampled through the use of a stratified method described below. Despite an effort to represent each age group equally, a shortage of available participants contributed to an uneven split in the three categories: 21 to 29 (N = 28), 30 to 39 (N = 29), and 40 and older (N = 23). Each participant was paid $25.00 for approximately three hours of participation.

Materials

Multiple Cue Probability Task

The MCPL task was based on one described by Rothstein (1986) and designed for public use by the Army Research Institute for the Social Sciences. The task was controlled by a computer program written in the Basic programming language for use on an IBM XT or AT style machine. Stimuli were displayed on a 26 x 19 cm Quadchrome video screen situated approximately 70 cm from subject's eyes.

The MCPL task has four vertical bars: Two "cue" bars

(A and B), one response bar manipulated by participants (S), and one criterion bar (C) representing the correct response. The participants' task on each trial was to predict the magnitude of C after observing the cues that were randomly selected by the computer program. Each trial had a 30-second time limit for the participants to manipulate their response (the S bar), and a 30-second study period to review the feedback provided after they made their prediction. The participant had the option to end the study periods by pushing the enter key. Thus, each trial lasted from a few seconds to one minute.

For each trial, participants viewed a standard screen. The width of "A", "B", and "S" remained constant at 13 mm. A space of 13 mm separated each bar. Eleven numbered horizontal lines on a vertical axis provided a reference scale from 0 to 10 for the height of each bar on the left side of the screen. Each cue had 19 possible values ranging from 1 to 10, with half steps between 1 and 10.

Each participant was presented with up to 100 trials in each of three levels of the task. If the participant reached the criterion (r_a equals .80) before the 100 trials were completed, the task ended at that trial. Each level corresponded to a particular level of task complexity. The three levels were distinguished by the mathematical relationship of cue bars to the criterion bar. In Level 1, the function was positive linear and is described in the following equation: $Y_c = .5X_A + .5X_B$, where X_A is the value for cue "A" and X_B the value for cue "B," and Y_c is the value of the criterion (note Y_c is sometimes referred to as Y_e or the environmental event in the Brunswick lens model literature). For example, if the "A" bar was 4 and the "B" bar was 8, the criterion ("C" bar) equaled 6.

In Level 2, called the "mixed treatment," the function for predicting the criterion was positive linear for one cue and inverted U-shaped for the other cue. The program randomly selects which cue (A or B) will have the inverted U-shape in this level. The inverted U-shaped function, g(x), was defined by substituting values of 1, 3, 5, 7, 9, 9, 7, 5, 3, 1, for the cue heights 1 to 10, respectively. The formula to compute the criterion was $Y = .5g(X_A) + .5(X_B)$ or $Y = .5(X_A) + .5g(X_B)$. For example, assume the "A" bar was selected by the computer as the curvilinear cue. If the "A" bar was equal to 9 (9 equals 3 when transformed by the function form $g(X_A)$) and the "B" bar was equal to 3, the criterion bar equaled 3.

In the third stage, the function form for predicting the criterion from the cues was inverted U-shaped for both cues, e.g., $Y = .5g(X_A) + .5g(X_B)$. Following the same example above, both the 9 and the 3 were transformed to 3 and 5, respectively, and the criterion was equal to 4.

In the present study, the nature of the MCPL task was deterministic, not probabilistic. Therefore, the r_e, the correlation between the cues and Y_c, is not relevant. Therefore, three of Dudycha and Naylor's four scores are associated with the MCPL task. First, the achievement index, r_a, is the correlation of the research participant's judgment of the criterion with the actual value of the criterion. Second, the linear predictability of the participant, r_s is computed on the basis of the regression of the participant's Y_s on the values of the two cues.

Third, the matching coefficient, r_m, is defined as the correlation between predicted Y_s and predicted Y_c.

Because Y_c is perfectly determined, r_m in this study is the correlation of Y'_s and Y_c.

Hypothesized covariates:

Career Path Appreciation Interview

The score from the Career Path Appreciation (CPA) interview was used as a measure of construing ability. The interview was an evaluation of the current scope of the person's capacity to cope with complex problem solving, sometimes referred to as cognitive complexity (Jacobs & Jaques, 1989). For illustrations of the cards used in this task, the phrases for the nine sets of phrase cards, and the protocol for administration including participant directions, see Appendix C.

There were two subtasks comprising the CPA as the career history portion was not used. In the first subtask, designed by Stamp in 1988, the participant was presented with a pack of symbol cards. Each card presented a number of symbols (one, two, or three) of a certain shape (circles, triangles, or squares) of one color (red, green, and blue), and one size (small, medium, or large). At the beginning of the task four display cards were set out, from left to right, in front of the participant. Three of the display cards illustrated various combinations of the four factors. One card contained two medium green squares, one card contained one small red triangle, and one card contained three large blue circles. The fourth card was blank. The participant was given a pack of 162 cards. On each card, there appeared a different combination of the four factors. For example, a card might have two large green triangles or three small blue circles. The participant's task was to discover and implement a rule for sorting the cards (by matching size and shape of the symbols on the card with size and shape of the symbols on the display card). During administration, the administrator provided the respondent with information about whether or not the placement was correct if the respondents placed a card in a nonblank pile. When a card was placed in the "blank card pile," no feedback was given. Additionally, participants were told that no previous trials could be reviewed.

Participants were given approximately 80 cards to discover the rule with no time restrictions. If ten consecutive cards were placed correctly, the card sort task was considered to be completed successfully. In general, if the rule was not discovered by the eightieth card, the administrator terminated the task and asked questions to discover what hypotheses the participant had tested. Although there are 162 cards in the pack, only half are usually necessary to estimate the respondent's thinking processes.

After the initial explicit directions, the participants controlled the task, and followed whichever personal problem solving strategies they preferred. This strategy was discovered through an interview that followed the symbol sort task. The purpose of the interview was to let participants disclose the way they thought about the task, the hypotheses they tested vis a vis the "sorting rule," and the strategy they used to solve the task.

Following the card sorting subtask, 54 phrase cards, arranged in nine sets of six cards, were given to the respondent one pack at a time. The respondent was asked to choose the card he or she felt "most comfortable with" and the card he or she felt "least comfortable with" for each set. In addition to selecting a "most" and "least" comfortable card, respondents explained in a sentence or two the rationale for their selections. Each card in a set of cards corresponds to the levels of complexity in the hierarchy proposed by stratified systems theory, discussed above. If respondents said that they felt most comfortable with a highly complex idea and least comfortable with a low complexity idea, they are assumed to operate at a higher level of complexity based on rules established by Stamp (1988) and Jaques (1989).

The experimenters trained in administration of the CPA tasks were Dr. Robert McIntyre, and graduate students Catherine Greenwald, Pamela Jordan, and Laura Hamill. Dr.

Thomas O. Jacobs, Army Research Institute, trained the four administrators. Dr. Jacobs has administered the CPA many times with Army personnel. The subjects' performance on the card tasks were videotaped. Because of the complexity in scoring, Dr. Jacobs scored the videotaped data. The scoring procedures were established by Gillian Stamp for administration and scoring. Dr Jacobs used these scoring procedures to assess a person's current level of construing. From this level, an anticipated mode is computed, which is an estimate of a person's cognitive functioning that takes into account (i.e., partials out) a person's age. In addition, a third index is computed based on the "most comfortable" card chosen for each of the nine sets of phrase cards in the second subtask. For each set, the respondent indicated one card of the six which they were "most comfortable" with, and then the rank of each card on the SST hierarchy is summed. This index, suggested by Jacobs (personal communication, 1991), and referred to as the phrase card average, is included in an attempt to explore a simple way of scoring the CPA.

Kirton Adaptation-Innovation Inventory

The KAI measures two styles of problem solving used by different personalities, the "adaptor" and the "innovator." The scores are on a continuum from 32 to 160, and as was pointed out above, neither style is considered more desirable than the other.

The KAI measures the proposed cognitive style defined by the adaptation-innovation theory. This theory proposes that problem solving, decision making, and creativity are closely interrelated concepts to which cognitive capacity (intelligence) is orthogonal.

Participants completed the 33 item questionnaire that has its own carbon scoring sheet attached on all four sides so that it is not visible to the participant. The total score for each participant, which indicates where on the continuum of adaptation-innovation he or she is, was used for this study. Norms provided in the KAI Manual indicate that sixty-seven percent of the people fall in the range from 80 to 112, with 96 as the median score. The low end of the scale (32 is the lowest possible score) indicates an adaptor type of creative style, and the high end of the scale, an extreme score of 160, indicates the innovator type of creative style.

<u>Mastermind</u>

Mastermind is a computerized game based on a board game by the same name, that involves the use of logic to discover the "secret code" involving the ordered placement of four "pegs." This discovery of a "secret code" is the identification of the relevant attributes of the conceptual rule. The four pegs can be any of the six colors, with the repetition of colors allowed (6^4 = 1,296 possible permutations). Thus, two possible "secret codes" could be green, green, red, yellow or green, red, yellow, blue. The term "peg" is used because the original board game involved placing pegs in holes, and getting feedback by means of black and white pegs.

In Mastermind, the participant must discover not only which colored pegs are included in the secret code, but in which order they belong. The feedback for each trial gives some information, and each trial is visible throughout the entire game. After the participant makes an initial guess about the four colors and their appropriate order, the feedback is provided. One black feedback peg, for example, indicates that one of the four chosen colors is correct and correctly placed. Two white feedback pegs would indicate that two of the chosen colors are correct, but incorrectly placed. Up to four feedback pegs can be earned, with the goal of receiving four black pegs.

A computer-generated random combination of pegs was used for each subject's practice trial. Each practice trial involved an explanation of the game rules and feedback rules, but no strategy advice from the experimenter. Following the practice trial, each participant played three additional games. One of those games was the six-peg color game, and the other two games were seven-peg color games. The number of trials used (up to ten) and whether or not the participant solved the secret code was recorded.

The scoring of Mastermind evolved over the course of the study. During the pilot phase, one or two practice trials of six pegs and a game of six pegs was thought sufficient to assess a participant's use of logic. The score was to have three elements: the amount of time used to finish the game, whether or not the game was solved, and the number of trials needed to solve the game (10 trials was the maximum possible) were thought to adequately assess performance.

The final scoring procedure, developed after the tenth research participant, measured whether or not the participant solved the puzzle and how many trials out of 10 he or she used to solve each puzzle. If the participant failed to solve the puzzle in ten trials, the computer game automatically ended, and the feedback for the last trial was scored using the scoring sheet that was developed. This scoring sheet assigned a number from five to twenty to a hierarchy of possible feedback pegs (see Appendix D).

Demographic Information

A demographic questionnaire was completed by the participant. The questionnaire requested the participants to indicate gender, age, graduate major, year of study, undergraduate degree(s), the number of quantitative courses the participant has completed, and the number of art courses completed (see Appendix E).

Graduate Record Exam

The Graduate Record Exam (GRE) General test assesses the skills that are essential to success in most graduate programs and is a generally accepted measure of scholastic ability. The General portion of the test consists of three sections: verbal, quantitative, and analytical. GRE scores for this study were obtained from the Office of the Registrar with the participants' written permission. Unfortunately, of the 80 participants, only 48 had GRE information on file with the Registrar.

Test Anxiety Profile

The Test Anxiety Profile (TAP) is a self-evaluation questionnaire that yielded 2 subscores: A Feelings of Anxiety score (FA) and a Thought Interference (TI) score. Each TAP inventory took approximately 3 minutes to complete. The inventory allowed measurement of the anxiety level of the participant in testing situations at three different times of the study: One was administered during the first session of the experiment, and the other two were administered at the beginning and end of the second session (see Appendix F). For the purposes of the present research, the two subscores (FA and TI) of each participant on the final anxiety questionnaire are of interest for the primary analyses.

Reliability data are available in the TAP Manual. Internal consistencies data ranged from .88 to .96, and test-retest reliabilities (done 7 weeks apart) ranged from .67 to .80. Validity data are described extensively in the Manual also. Convergent and discriminant validities were examined. Sarason's Test Anxiety measure correlated with combined measures of the TAP to provide convergent validity (Oettig & Deffenbacher, 1980). Unique clusters of anxieties from cluster analysis of the different scales of the TAP showed discriminant validity. The FA items are taken from another measure that has been shown to have validity and reliability through a series of studies spanning 12 years. Construct validity is thus supported for measures of concept-specific anxiety, or anxiety that is specifically measured by each of the concept categories the TAP can measure such as math test anxiety or anxiety due to presenting a speech.

Procedure

Participants were randomly selected from a master list of graduate students in the College of Arts and Letters, the College of Education, the College of Engineering, the College of Health Sciences, and the College of Sciences. The stratified sample strategy selected participants by age group within each college. Age groups established were 21 to 29 years old, 30 to 39 years old, and 40 and older. Five people per age group per college were randomly selected to generate a list of 90 possible participants. Each person selected was then called by one of the experimenters and asked to participate following the protocol established (see Appendix A). Eighty-five percent of the originally selected people were unavailable, so a new list of potential participants was generated. Several of these lists were exhausted in the attempt to recruit participants. The recruiting calls consisted of the caller's introducing herself and the purpose of the call, and then the study was outlined briefly. If the person indicated an interest in participating, a time was scheduled for the first session. A protocol of a sample conversation is in Appendix A.

The experiment consisted of two sessions. Participants were interviewed individually. The first session lasted between one hour and one and one half hours. The second session lasted approximately two hours. In the first session, the participants read and signed a standard informed consent form and an additional form (see Appendix B) on which they confirmed their understanding that payment was contingent on completing both sessions, and gave permission for the release of the participant's GRE scores from the university administration.

The orientation period took about 5 minutes. After this, each participant was given the demographic questionnaire (10 minutes), followed by the Anxiety questionnaire (5 minutes), and the Kirton Adaptation-Innovation Inventory (10 minutes).

The Career Path Appreciation interview was the next task for the participant in Session 1. The experimenter explained the instructions for the CPA Phrase Card and Card Sort subtasks, and the need for the use of the videocamera for scoring purposes. The video camera was set up above the test table to record the nonverbal and verbal cues that the participant exhibits. The CPA card tasks were administered by the trained experimenters and the videotapes were sent to Dr. T. O. Jacobs for final scoring and evaluation. Assigned participant numbers were used on all materials to preserve confidentiality.

After completion of the CPA tasks (30 to 45 minutes), each participant completed the Mastermind computer task (one practice trial and three experimental trials) in 20 to 30 minutes.

The second session began with the participant's completing the Anxiety questionnaire (3 minutes) for the second of three times. The participant was told the instructions for the task both verbally and on the computer screen. Once the participant understood the instructions and asked questions, the task began. When the participant completed Level 2 (of the three levels), he or she was invited to take a 15 minute break. After the break, the participant continued the task to complete Level 3. Upon completion of the MCPL task, the participant completed the Anxiety questionnaire for the final time. When all tasks were completed, the formulas for the three levels of the multiple cue probability task were explained to the participants. Finally, all participants' questions were answered, each person was thanked for their participation and paid $25.00.

Participants generally completed the two sessions within 7 days, with at least 24 hours between sessions. However, two participants completed all tasks in one session because of personal time constraints.

III. RESULTS

Distributions of Variables

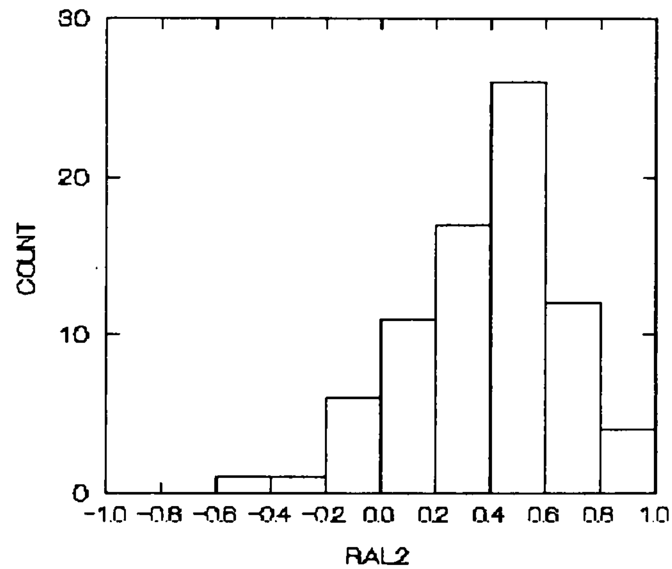
Frequency distributions for variables other than the multiple cue probability task levels are presented in Appendix G. The most important finding from these figures, presented here, is that the distributions of Level 2 and Level 3 for the achievement index (r_a) and the matching index (r_m) do not follow the same pattern.

Achievement Index

The achievement index for Level 2 has a mean of .381 (SD = .278) and a median of .416 (see Figure 2) while the achievement index for Level 3 (Figure 3) has a lower mean (M = .198, SD = .276) and median (.180). This indicates lower achievement on Level 3 than Level 2. In order to test that observation, a repeated measures analysis of the means for Level 2 r_a and Level 3 r_a show a significant difference, $F(1, 159) = 20.36$, $p < .01$.

Matching Index

The distribution of the matching index on Level 2 (Figure 4) has a mean of .519 (SD = .384, median = .589) and the distribution of the matching index for Level 3 (Figure 5) shows a lower mean (M = .321, SD = .473, median = .371). This indicates many participants were able to match their strategy to the function rather well in the second level,



<u>Figure 2</u>. Multiple cue probability learning task Level 2: A frequency distribution of participants' achievement indices.

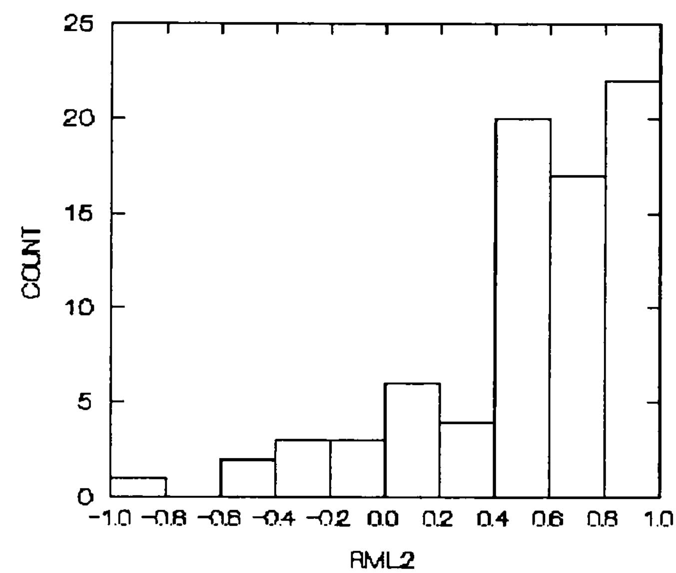
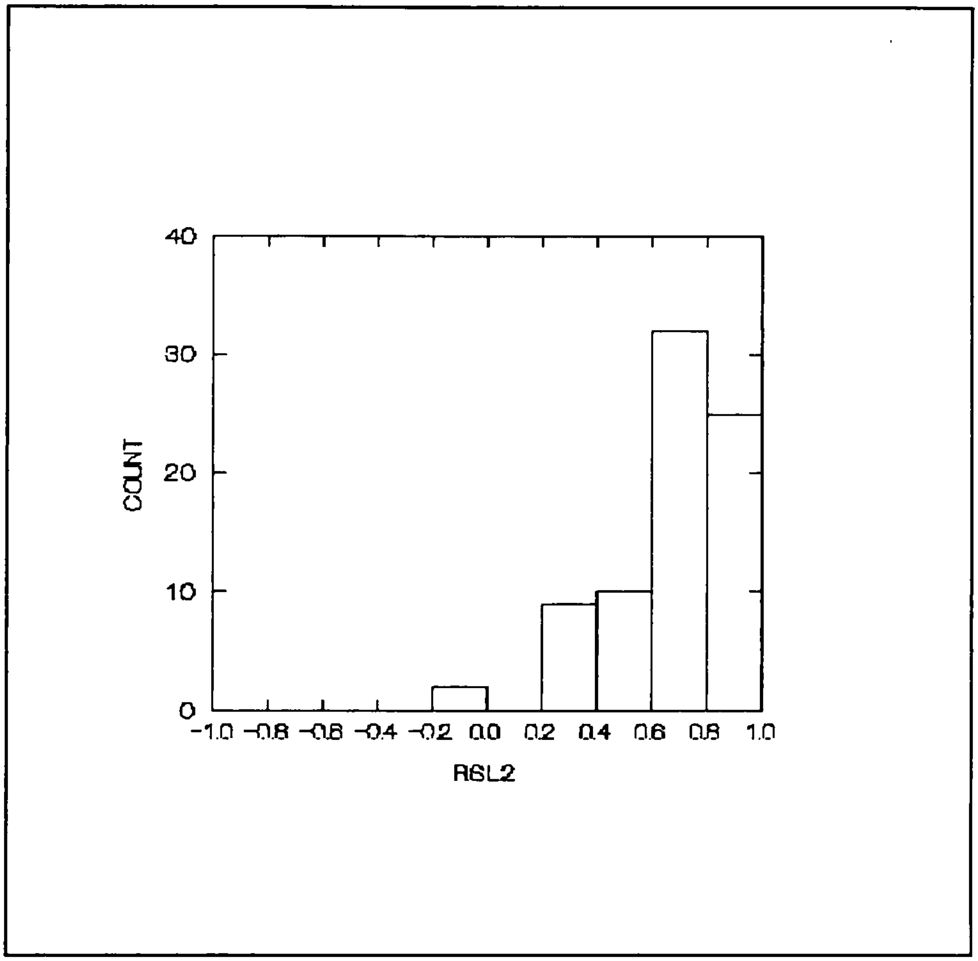


Figure 3. Multiple cue probability learning task Level 3: A frequency distribution of participants' achievement indices.



<u>Figure 4</u>. Multiple cue probability learning task Level 2: A frequency distribution of participants' matching indices.

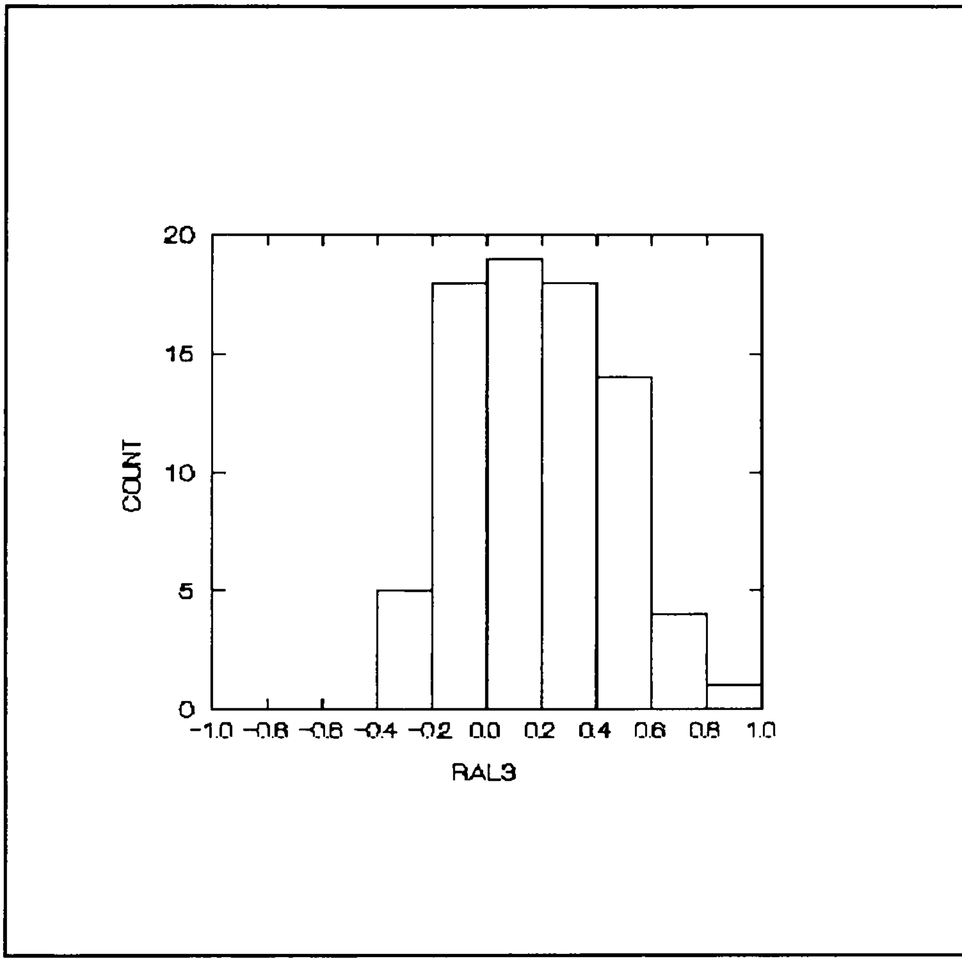


Figure 5. Multiple cue probability learning task Level 3: A frequency distribution of participants' matching indices.

but not the third. This was also tested with a repeated measures analysis of the means. Level 2 r_m and Level 3 r_m are significantly different ($F(1, 159) = 10.91$, $p < .01$).

Linear Stability Index

The linear stability index distribution for Level 2 (Figure 6) has a mean of .671 (SD = .215, median = .723) and for Level 3, the stability index distribution (Figure 7) has a mean of .554 (SD = .183, median = .571). There is a significant difference among r_s for Level 2 and Level 3 ($F(1, 159) = 18.22$, $p < .01$).

Analyses

Correlations among the predictor variables such as the Career Path Analysis interview, the Kirton Adaptation-Innovation Inventory, the Mastermind logic game, demographic information such as gender, age, and college, the GRE scores, and the Test Anxiety Profile (TAP) were used to test the hypotheses.

Pearson Product Moment correlation analyses were conducted for the three MCPL criterion indices, r_a, r_m, and r_s, with all predictors. The reader will recall that the achievement index (r_a) is a measure of degree of agreement between the criterion values and responses of the participant over 100 trials. The matching coefficient (r_m) indicates the degree to which the participant's strategy or policy matches the optimal criterion equation. The index of subject consistency (r_s, or equivalently, R_s). The index of subject consistency measures the degree to which the

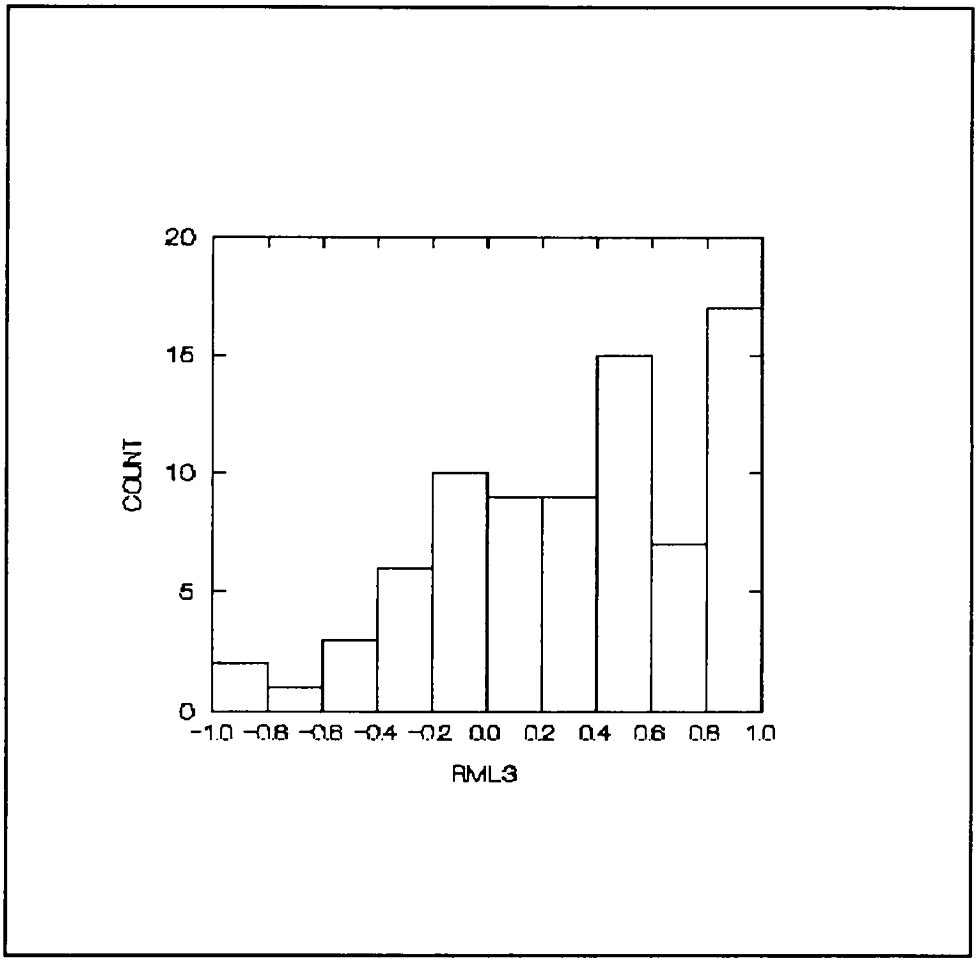


Figure 6. Multiple cue probability learning task Level 2: A frequency distribution of participants' linear stability indices.

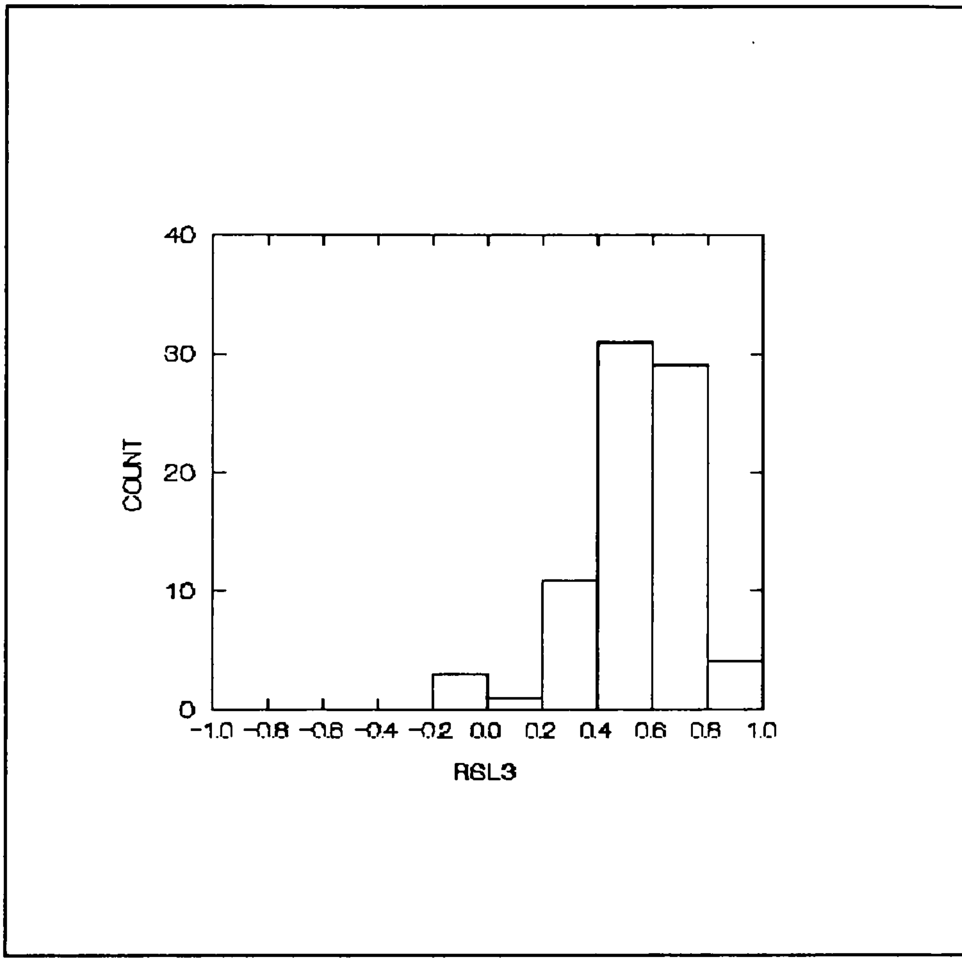


Figure 7. Multiple cue probability learning task Level 3: A frequency distribution of participants' linear stability indices.

participant consistently uses his or her strategy across the trials. These criterion indices were calculated for the last 25 trials of the second and third level of the MCPL task. It was assumed that a decision maker's learning curve in the MCPL task reached asymptote in the last 25 trials of the task. In other words, it was assumed that a decision maker's policy was formed and stabilized after the first 75 trials of the MCPL task and that an adequate representation of his or her performance can be captured by computing r_a, r_m, and r_s in the last 25 trials. Indices for the first level of the MCPL task were not calculated because of the lack of variance in performance. That is, most participants solved the first level within 15 trials.

The correlations between the predictor variables and the six MCPL indices are summarized here and presented in Table 3. There were no significant correlations between the current level and mode (the CPA measures) and the MCPL indices. An unanticipated finding was with the third index associated with CPA performance suggested by T. O. Jacobs (personal communication, August 10, 1991). The CPA phrase card average index is the mean of the values associated with the nine "most comfortable" phrase cards that the respondent chose. This measure correlated positively with Level 2 r_a (r = .301, $t(76)$ = 2.752, $p < .05$) and Level 2 r_m (r =.330, $t(76)$ = 3.048, $p < .05$). This indicates that the higher the level of abstraction of work described on the phrase cards, the higher the achievement and matching strategies for the

Table 3

	r_a		r_m		r_s	
	MCPL LEVEL					
Predictor	2	3	2	3	2	3
CPA:						
Current	.045	-.129	.113	-.062	-.031	-.171
Average	.301*	.036	.330*	.094	.133	-.079
Mode	-.018	-.148	.023	-.090	-.104	-.197
KAI	.209	-.021	.314*	-.005	-.138	-.129
Mastermind	.035	.065	.016	.086	.032	-.066
Demographics:						
Quant Classes	-.136	.140	-.096	.178	-.026	-.160
Art Classes	.096	-.088	.058	-.086	.168	.081
Art Hours per week	.158	-.155	.151	-.117	.059	-.123
Age	.178	.065	.218	.115	.129	-.035
Gender	.266*	.001	.292*	.013	.063	.046
GRE:						
Verbal	-.004	.070	-.031	.092	.109	-.036
Quantitative	.310*	.255	.236	.258	.198	.054
Analytical	-.120	.210	-.141	.217	.021	-.020
Anxiety Time 3:						
Feelings	-.221	-.110	-.275*	-.091	-.106	-.071
Thoughts	-.232*	-.179	-.246*	-.170	-.191	-.152

* $p < .05$

Level 2 MCPL task.

The correlations among KAI and the MCPL indices were corrected for attenuation (unreliability). This was done in order to estimate the true correlation of the theoretical construct creativity from the imperfect observed measure (the KAI) of creativity (Cohen, 1975). This is done by dividing the observed correlation by the square root of the product of the reliabilities of the measures. Since the MCPL task had perfect reliability, the correlations among KAI and MCPL indices were divided only by the square root of the reliability of the KAI. The correlation of the KAI and the r_m of Level 2 was significant (r = .314, $t(76)$ = 2.883, $p < .05$), which indicates that the more a person showed a tendency to be an innovator, the better his or her matching strategy for the second MCPL task. There were no significant correlations among the KAI and the other MCPL indices.

There were no significant correlations among the Mastermind average score and the MCPL indices.

For the demographic data, significant correlations were found among gender and the Level 2 indices r_a and r_m. The correlation of gender with r_a was .266 ($t(76)$ = 2.406, $p < .05$), and indicated that men tended to have significantly greater achievement indices than women. The correlation of gender with r_m was .292 ($t(76)$ = 2.662, $p < .05$), and indicated men tended to have significantly greater matching coefficients than women. In order to identify if the relationship was moderated by Quantitative GRE, KAI, and anxiety (TI), post hoc partial correlations were calculated. These three variables were partialed because of their significant correlations with gender (see Appendix H). The partial correlations indicated that the relationship among gender and the MCPL indices was moderated by the Quantitative GRE (corrected for attenuation), KAI (corrected for attenuation), and anxiety (TI) scores. Thus, there were no real differences between men and women on the MCPL indices.

There was one significant correlation among the GRE scores and the MCPL indices. After correcting for attenuation (unreliability), the correlation of Quantitative GRE and Level 2 r_a became significant ($\underline{r}$ = .310, $\underline{t}(44)$ = 2.163, $\underline{p} < .05$). This indicates that participants with higher Quantitative GRE's scores had higher achievement on Level 2.

Significant correlations were found among the anxiety measure taken after the Level 3 MCPL task and the MCPL indices for Level 2. The feelings of anxiety measure was significant with Level 2 r_m ($\underline{r}$ = -.275, $\underline{t}(76)$ = -2.397, $\underline{p} < .05$), which indicates increasing accuracy of matching in the second MCPL task with decreasing anxiety measured after the third MCPL task. Similarly, the thought interference aspect of anxiety was significant with Level 2 r_a ($\underline{r}$ = -.232, $\underline{t}(76)$ = -2.079, $\underline{p} < .05$) and Level 2 r_m ($\underline{r}$ = -.246, $\underline{t}(76)$ = -2.213, $\underline{p} < .05$). This showed increasing performance on MCPL (matching and achievement) with decreasing thought interference.

Additional auxiliary findings that were not hypothesized are presented in Appendix H.

Secondary multiple correlation analyses were conducted post hoc to identify patterns of variables in their relation to the MCPL indices. No variables contributed strikingly to the variance in any meaningful way.

IV. DISCUSSION

Multiple cue probability learning research is extensive, but no efforts have been made to investigate the contributions of individual differences to MCPL performance. In an effort to begin research in this area, this study has assessed the correlation between creativity, construing ability, anxiety, verbal ability, quantitative ability, problem solving ability, and MCPL performance. MCPL performance on a computer task was measured in three ways: an achievement index, a matching coefficient, and a stability coefficient (r_a, r_m, and r_s) were computed for both Level 2 and Level 3.

The first hypothesis was that there would be a positive correlation between scores on the Career Path Appreciation interview, which measures construing ability and cognitive complexity of problem solving strategies, and the different MCPL performance indices. At first glance, this hypothesis seemed to be unsupported. However, there were significant correlations between the CPA phrase card average index and each of the MCPL indices. This suggests partial support for the hypothesis.

In order to understand these findings, the relationship of other variables with CPA phrase card average was investigated. This investigation indicated that CPA phrase card average correlated with other measures as might be expected, such as creativity (KAI). Consequently, it seems reasonable to conclude that the correlation between CPA phrase card average and MCPL were not Type 1 errors. There are several explanations for the finding. The primary CPA scores, CPA current level and CPA mode, are composite judgments based on the assessment of the two tasks comprising the CPA interview. It may be the case that the composite judgments are unreliable. This seems unlikely given respectable performance of the CPA in previous research (Jacobs & Jaques, 1989).

Alternatively, the two tasks comprising the CPA interview in this study may represent two distinguishable factors. For example, the phrase card sort task on which the CPA phrase card average score is based may pertain to some index of the respondents' self-awareness of their ability to tolerate ambiguity. The symbol sort may focus more on the cognitive capabilities of dealing with complex problems. When treated by the scorer as a single factor, the factorially complex composite score that results may not be ideally suited to correlate with the MCPL measures. This should be addressed in follow-up research.

The next hypothesis was that there would be a positive correlation between the MCPL task indices and the KAI score. This was partially supported. The correlation of the KAI and the r_m of Level 2 was significant. There were no correlations among the KAI and the other MCPL indices. The correlation indicates that the people who scored more toward the "innovator" side of the KAI continuum for creativity had better matching strategies on the second level of the MCPL task. It appears as though "innovator" participants are better at inferring the nature of the function for a "mixed" MCPL task (Level 2) than the adaptor. It is not clear why this finding does not extend to the Level 3 "completely nonlinear" MCPL task. The task may be too difficult for most participants, a fact that is suggested by the similarity of the two levels on all three MCPL indices. Alternatively, moderators such as decreasing motivation due to fatigue or boredom may have contributed to an attenuation of the correlation between KAI and MCPL indices in Level 3.

The next hypothesis was that there would be a positive correlation between concept-learning problem solving performance on "Mastermind" and the MCPL task indices. It was not supported. There were no significant correlations among the Mastermind average score and the MCPL indices. There are two reasons for the lack of significant relationship besides the possibility that no relationship exists. The first pertains to scoring. The scoring procedure was not developed at the time of the creation of the computer version of the game. It was instead developed by the researchers during the study as new problems and information became apparent. This lack of reliability in the scoring method could have contributed to a lack of findings. The criteria chosen to represent Mastermind performance (number of trials to discovery of the secret code) may not be an appropriate measure of Mastermind performance. Laughlin, Lange, and Adamopoulos (1982) provide a discussion of two types of strategies, focusing and tactical, that were used in their study of a simplified version of Mastermind. It is possible that their measurement of strategies, as opposed to achievement, in Mastermind is a better measure of concept-learning problem solving.

The second pertains to the fact that each participant played a different set of Mastermind games. This was due to the way the computerized game generated random game data. At first, the random generation of computer games for each person seemed useful. However, further investigations suggest that Mastermind games vary with respect to their difficulty, a factor not controlled in this study. The difficulty of the Mastermind game varies as a function of the combination of pegs. Some participants had "easy" games while others had difficult games. An "easy" game could be explained as a participant guessing "red, green, black, blue" on the first trial, when the correct answer was "red, green, black, black," which would give the feedback that the participant's random guess for the first trial was 75% correct, with only one more colored peg to discover in nine trials. A more difficult game could be explained as one where the correct combination is "red, green, green, green" and the participant tries several possible color combinations (perhaps using all ten trials) before realizing that the combination had only two colors, with one repeated three times. The problem of variable game difficulty was dealt with by having each participant play three games, and the average score was used as the measure.

Due to the possibility that the lack of findings could be related to these two moderators, future research on this variable should not be ruled out. Instead, a controlled test of Mastermind with the scoring hierarchy developed by the researchers and the same combination of pegs for every participant' game should identify a more valuable correlation with MCPL indices. The strategies described by Laughlin, et al. might also be useful in developing scoring hierarchies for future research.

One last comment about Mastermind is in order. Despite the two confounding elements just described, Mastermind scores were apparently sufficiently reliable to correlate with other measures that one would expect. In particular, Mastermind correlated positively with Verbal GRE, Quantitative GRE, and Analytical GRE scores. These significant relationships may suggest that the lack of relationship between MCPL and Mastermind represents not a Type I error, but an accurate depiction of the state of affairs. Such can be confirmed with an improvement in the scoring and game procedures.

The next hypothesis was that experience level with mathematical problem solving would correlate positively with MCPL performance. This was not supported. Perhaps the type of quantitative experience reported (number of classes) had little to do with MCPL performance. It is also possible that the questions pertaining to information about former quantitative classes were not defined well enough to elicit the same kind of information from every participant. It was difficult for many people to remember how many quantitative classes they had taken, and there was no method for the experimenter to evaluate the level of difficulty of the class or the actual amount of quantitative problem solving required in each class. In future research, perhaps asking for transcripts would be an effective way of getting a reliable estimate of quantitative classes.

Unanticipated Effects

The effects of gender on MCPL performance have not been reported in the current literature, but the present investigation shows that there were significant correlations between gender and the Level 2 indices r_a and r_m. The correlation of gender with r_a indicated that men had a significantly greater achievement index than women. The correlation of gender with r_m indicated men had a significantly larger matching coefficient than women. That this is not repeated for Level 3 is not understood. It is interesting to note a pattern of relationships between gender and other variables such as Quantitative GRE scores, the KAI, and anxiety (TI) which indicates that men tended to have the following: (a) greater Quantitative GRE scores; (b) scores on the KAI that represent the "innovator" end of the continuum more than women; and (c) less anxiety (TI) than women.

In order to assess the effect that other predictors had on the relationship between the MCPL indices and gender, post hoc partial correlations were computed. Since the men's scores on the KAI, the Quantitative GRE scores, and the thought interference test anxiety scores were significantly different from the women's, those predictors were partialed out of the correlation among gender and Level 2 r_a and r_m. This decision was made based on Hyde's (1990) explanation that reviews of gender differences show men have greater visual-spatial ability and greater mathematical ability than women. Whether or not this is due to genetics or socialization is still not clear, but the differences are there. When the partial correlation was calculated, the relationship disappeared. Since the gender difference disappeared in the present study when the Quantitative GRE scores were partialed out, the conclusion that men perform better on MCPL tasks seems unwarranted. Instead, people who have higher Quantitative GRE scores perform better on MCPL tasks. In this study, the higher Quantitative GRE scores belonged to the men.

The exploratory hypothesis that age would show a positive correlation with the MCPL indices was not supported. However, the hypothesis that age would show a positive correlation with MCPL indices was based on the logic that older participants perform better on the CPA interview. Since most of the CPA measures did not show a significant correlation with MCPL measures, it is not surprising that the relationship between age and MCPL indices was not significant.

The hypotheses that the Verbal GRE score and the Analytical GRE score would positively correlate with MCPL performance were not supported. With a correction for attenuation, the relationship between Level 2 r_a and Quantitative GRE was significant. This is mentioned because the tests of the relationship between GRE scores and MCPL performance suffered from reduced power due to the fact that GRE scores were only available for 48 of the 80 participants. This fact had not been anticipated prior to the study. In fact, only graduate programs that required GRE's were sampled. Thus, future research similar to this study should continue to investigate GRE scores as a predictor of MCPL performance in order to obtain more information.

Anxiety about performance was hypothesized to be negatively related to MCPL performance. The anxiety measure had two subscores: Feelings of Anxiety (FA) and Thought Interference (TI). The feelings of anxiety measure was significantly correlated with r_a and r_m from Level 2. This indicates decreasing achievement at Level 2 with increasing anxiety measured post-MCPL Level 3. Similarly, the thought interference aspect of anxiety correlated significantly with r_a and r_m from Level 2. Although it is difficult to confirm with the present data, it appears reasonable that participants' anxiety-related feelings and thoughts rose as a result of the Level 2 task and remained stable through

Level 3. This would explain the relationship between Level 2 performance on the MCPL task and the anxiety-related measures taken after Level 3. It is assumed that the distribution of Level 3 performance indices were not conducive to correlations with the anxiety measures, a fact that is suggested in the lack of correlation between any of the hypothesized correlates and Level 3 MCPL measures. In future research, it is suggested that anxiety measures be administered after each Level of MCPL in order to understand more precisely the nature of the relationship between anxiety at each level of MCPL and MCPL performance.

Limitations of the Research

Missing Data

This study had limitations. Missing GRE scores for 40% of the participants was an unanticipated problem in this study. Not every program who listed GRE's as a requirement in the university application (our criterion for choosing majors to include in the experiment) actually required the students to have them or to provide them to the permanent file.

Additionally, due to a flaw in the design of the procedure, an anxiety measure was not administered after the Level 2 MCPL task. This led to limitations in the interpretation of the findings among anxiety and the MCPL indices. The anxiety measures administered prior to the MCPL task could not be expected to reflect anxiety felt about the task, and the measure administered after Level 3 may not adequately reflect anxiety participants felt during

Level 2. The frustration that the participants was very obvious to the experimenters. Many participants made comments about their frustration, and some even expressed anger that they could not solve the puzzles. Other participants felt anxiety because they had never participated in a psychology experiment, and they asked about their progress. Thus, future research should include an anxiety measure after both difficult MCPL levels.

Level 3 Limitations

There are three reasons why Level 3 MCPL performance may not have correlated with the predictors that correlated with Level 2 performance. First, fatigue effects, boredom, or lack of motivation effects were identified with anecdotal evidence collected during the post-MCPL interview: Many participants mentioned just wanting to "get it over with" or that they thought it was boring and tiresome. Second, cognitive set effects are suggested because learning the Level 2 function first may have created an inability to learn the Level 3 function. And third, the participants' Level 3 performance was found to be significantly less successful than their Level 2 performance, which could indicate that Level 3 was too difficult or that there may have been an insufficient number of trials for learning the function of Level 3. Dudycha and Naylor (1966) gave participants 200 trials to learn the nonlinear function and found that, across blocks of 50 trials, achievement reached a peak at 100 trials, and then leveled off for the last 100 trials. If learning is still taking place up to 100 trials, as Dudycha and Naylor suggest, the peak performance may not have been captured in the present study's calculated indices.

Therefore, the following modifications are recommended. It would be useful to counterbalance the second and third levels of MCPL tasks in order to rule out fatigue effects and cognitive set effects. Alternatively, groups could be compared by matching participants across other variables (gender, age, college) and assigning one group to perform Level 2 and the other group to perform Level 3. And additional trials could be added to the Level 3 task in order to identify if 100 trials is really sufficient or not.

Mastermind Limitations

Limitations were found with the Mastermind game, but excluding Mastermind from future research in the measurement of concept-learning problem solving is not recommended. Several improvements in the use of Mastermind as an experimental tool are needed. First, refined scoring procedures should be established based on the hierarchy presented in this research. Second, standardized games should be used. The computer program could be updated to provide a menu of specific standard games and to allow the storage of performance data. Thus, the experimenter could choose games so that all participants could be compared on the same game. The game, in its computerized form (after appropriate modifications), does provide the possibility of assessing a participant's problem solving skill with a minimum of scoring bias on the part of the experimenter.

Development of computerized scoring methods and data storage would also facilitate lack of experimenter error and ease of data collection. In addition, participants like the game, making data collection quite easy. Several participants were excited about playing Mastermind and even asked if they could play an extra game or two!

Anecdotes and Observations

It was interesting to discuss with participants what they thought about each experimental task and how they chose a solution. Participants were asked what types of strategies they tried after Level 2 and after Level 3 of the MCPL task and their answers were recorded. Three people mentioned that, for the MCPL task, they did not look at the numbers of the bars in order to do mathematical formulas at all. Instead, they chose to press through several trials in a few seconds (without choosing an answer) just to get a feel for the "pattern" of the bars. At least one of these people solved the task. Similarly, people used other spatial cues such as "if the A bar is larger than the B bar, then the answer is larger (greater than 6)" or "if the space (difference) between the two bars is small, and the bars are tall, then the answer is large." Others mentioned linear strategies they used (i.e., 10 minus the large bar's value plus the small bar's value) that were actually successful in aiding a solution to nonlinear MCPL tasks. Although there were a few commonalities between participant's initial simple attempts at solution (i.e., try the average of the bars, or add the difference ($|A-B|$) to the smaller bar), it seemed there were almost as many different types of solutions as there were participants. The answers to the question "what strategies were you trying" are not in any quantifiable form, but the information possibilities are intriguing enough to point to the need to discover patterns of individual strategies that are successful and unsuccessful. This could be accomplished in future research through a more structured post-MCPL questionnaire with a developed MCPL-strategy scoring system.

Conclusions

Is MCPL performance an individual difference variable? This was the question presented at the beginning, but the answer is not a simple, straightforward one. There is some evidence that MCPL performance is an individual differences variable, but that evidence is not clear. Several predictors (CPA phrase card average, anxiety, creativity (KAI), and gender) covaried with MCPL Level 2 indices. There were "borderline" relationships with GRE scores that deserve more attention than the present study could provide with the lack of available data. Despite the limitations of this exploratory study, the results begin to suggest that MCPL performance is a function of a person's style of problem solving, quantitative intelligence, resistance to frustration and anxiety, and tolerance for ambiguity. This suggests the beginning stages of a theory of MCPL performance. Consequently, this research is successful in that it points to specific and refined research designed to improve understanding of multiple cue probability learning.

References

Athey, T. R., & McIntyre, R. M. (1987). Effect of rater training on rater accuracy: Levels-of-processing theory and social facilitation theory perspectives. Journal of Applied Psychology, 72, 567-572.

Brehmer, B. (1972). Cue utilization and cue consistency in multiple-cue probability learning. Organizational Behavior and Human Performance, 8, 286-296.

Brehmer, B. (1973). Note on the relation between single-cue probability learning and multiple-cue probability learning. Organizational Behavior and Human Performance, 9, 246-252.

Brehmer, B. (1973). Effects of task predictability and cue validity on interpersonal learning of inference tasks involving both linear and nonlinear relations. Organizational Behavior and Human Performance, 10, 24-46.

Brehmer, B. (1987). Note on subjects' hypotheses in multiple-cue probability learning. Organizational Behavior and Human Decision Processes, 40, 323-329.

Brehmer, B., Alm, H., & Warg, L. E. (1985). Learning and hypothesis testing in probabalistic inference tasks. Scandinavian Journal of Psychology, 26, 305-313.

Cohen, J., & Cohen, P. (1975). Applied Multiple Regression/ Correlation Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Dudycha, L. W., & Naylor, J. C. (1966). Characteristics of the human inference process in complex choice behavior situations. Organizational Behavior and Human Performance, 1, 110-128.

Hammond, K. R., Summers, D. A., & Deane, D. H. (1973). Negative effects of outcome-feedback in multiple-cue probability learning. Organizational Behavior and Human Performance, 9, 30-34.

Hyde, J. S. (1990). Meta-analysis and the psychology of gender differences. Signs: Journal of Women in Culture and Society, 16(1), 55-73.

Jacobs, T. O. (August 10, 1991) Personal communication.

Jacobs, T. O., & Jaques, E. (1989). Leadership in complex systems. In J. A. Zeidner (Ed.), Human Productivity Enhancement. Volume II. Organizations and Personnel. New York: Praeger Publishers.

Jacobs, T. O., & Jaques, E. (1990). Military executive leadership. In K. E. Clark & M. B. Clark (Eds.), Measures of Leadership (pp. 281-295). New Jersey: Leadership Library of America.

Kelly, G. A. (1980). A psychology of the optimal man. In A. W. Landfield and L. M. Leitner (Eds.), Personal Construct Psychology, Psychotherapy and Personality, (pp. 18-35). New York: John Wiley & Sons.

Klayman, J. (1988). Cue discovery in probabilistic environments: Uncertainty & experimentation. Journal of General Psychology: Learning, Memory, and Cognition, 14(2), 317-330.

Kirton, M. J. (1987). Kirton adaptation-innovation inventory manual, 2nd ed. Hatfield, U.K.: Occupational Research Centre.

Landfield, A. W., & Leitner, L. M. (1980). Personal construct psychology. In A. W. Landfield and L. M. Leitner (Eds.), Personal Construct Psychology, Psychotherapy and Personality (pp. 4-12). New York: John Wiley & Sons.

Laughlin, P. R., Lange, R., & Adamopoulos, J. (1982). Selection strategies for "Mastermind" problems. Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 475-483.

Meehan, A. M., & Overton, W. F. (1986). Gender differences in expectancies for success and performance on Piagetian spatial tasks. Merrill-Palmer Quarterly, 32(4), 427-441.

Nystedt, L. & Magnusson, D. (1982). Construction of experience: The Construction Corollary. In J. C. Mancuso and J. R. Adams-Webber (Eds.), The Construing Person (pp. 35-44). New York: Praeger Publishers.

Oettig, E. R., & Deffenbacher, J. L. (1980). Test anxiety profile manual. Ft. Collins, CO: RMBSI.

Rock, D. A., Werts, C., & Grandy, J. (1982). Construct validity of the GRE Aptitude test across populations--An empirical confirmatory study. (GRE Board Professional Report GREB No. 78-1P; ETS Research Report 82-57). Princeton, NJ: Educational Testing Service.

Rothstein, H. G. (1986). The effects of time pressure on judgment in multiple-cue probability learning. Organizational Behavior and Human Decision Processes, 37, 83-92.

Slovic, P., Fischhoff, B., & Lichtenstein, S. Behavioral Decision Theory. In M. R. Rosenzweig & L. W. Porter (Eds.), Annual Review of Psychology, (Vol. 28). Palo Alto, CA: Annual Reviews, 1977.

Sniezek, J. A., & Naylor, J. C. (1978). Cue measurement scale and functional hypothesis testing in cue probability learning. Organizational Behavior and Human Processes, 22, 366-374.

Stamp, G. P. (1988). Longitudinal research into methods of assessing managerial potential (ARI Technical Report 819). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (DAJ45-86-C-0009).

Steinman, D. O. (1976). The effects of cognitive feedback and task complexity in multiple-cue probability learning. Organizational Behavior and Human Performance, 15, 168-179.

Van Bolhuis-Bouma, A., & Oostlander, A. M. (1972). Multiple-cue probability learning: An experiment with circular dimensions. Organizational Behavior and Human Processes, 7, 383-396.

Wiggins, J. S. (1973). Personality and prediction. Principles of Personality Assessment. Reading, MA: Addison-Wesley Publishing Company, Inc. 156-165.

Appendix A

PROTOCOL FOR RECRUITING PHONE CALLS

"Hello, may I speak with _______________. Hi. My name is Cathy Greenwald and I am a graduate student at Old Dominion. I am working on my Master's thesis in Psychology and I am calling other graduate students and asking them to help me by participating in my study. Would you be interested?"

At this point, the contacted person either says he or she is too busy or is uninterested or asks for more information. If they are interested, I continue:

"Well, I am studying how people make decisions and use logic to solve problems. I am paying $25 for approximately 3 hours of work that is split into two sessions. The first session generally takes from one hour to an hour and a half, and it involves questionnaires, two card tasks, and a computer game. The second session is between one and a half to two hours and is another computer task. Would you like to schedule a time for your first session?"

At this point, the person either declines to participate, or agrees on a time to meet the experimenter at a designated place.

Appendix B

INFORMED CONSENT FORM CONTINUED

Below are additional agreements we need for your participation in this study. Please read them carefully as the experimenter reads them aloud to you. If you understand the statements and agree to participate, please sign your name.

I, ____________________________, agree to the release of my Graduate Record Exam (GRE) scores from the university Office of the Registrar to the experimenters only. I understand that my scores will not be released to any other person, or used for any purpose other than the present research. I understand that these scores will be kept confidential.

Signed ____________________________

I, ____________________________, agree to the following payment plan: I will receive $5 per hour for every hour I participate (approximately 5 hours, minimum payment $25) ONLY when I finish the entire experiment. I understand that I may leave at any time, but I will be paid if and only if I complete the experiment.

Signed ____________________________

WITNESS____________________________ Date______________

Appendix C

Career Path Appreciation

Sample Phrase Cards:

Write in "M" next to most comfortable card and "L" next to least comfortable card.

Set	A -	General principles
___	I	Work to a complete set of instructions
___	II	Work within a given framework
___	III	Work with connections even if particular links unclear
___	IV	Work in abstracts and concepts
___	V	Work with a minimum of preconceptions
___	VI	Define the horizons of the work

Set	B -	General process
___	I	Do one thing at a time
___	II	Focus on one part of the task at a time
___	III	Co-ordinate by drawing together many separate strands
___	IV	Compare the merits of alternative options
___	V	Establish new relationships between unrelated material
___	VI	Use words, ideas and theories as tools

Set	C -	Overall approach and attitude to rules
___	I	Follow the rules
___	II	Work within the rules
___	III	Extrapolate from the given rules
___	IV	Look for the pattern of the rules
___	V	See the rules as guides to action
___	VI	Redefine the rules

Set	D -	Approach to each problem
___	I	Follow instructions carefully
___	II	Approach each task in own right
___	III	Take a systematic approach
___	IV	Span a broad spectrum and also focus in detail
___	V	Restructure the task
___	VI	Transcend the task

Set	E -	First action on the problem
___	I	Do first things first
___	II	Break up a problem into separate parts
___	III	Look for sequences or common relationships
___	IV	Analyze problems by searching for underlying structure
___	V	Create an overall picture of the problem
___	VI	Consider the context of the problem

Set	F -	Reaction to uncertainty
___	I	Use your common sense
___	II	Allot a specific amount of time to each task
___	III	Tolerate uncertainty
___	IV	Handle ambiguity by developing opposing point of view
___	V	Expect task will be transformed while it's in progress
___	VI	Transform the task to create uncertainty

Set	G -	Reaction to "gaps" in knowledge
___	I	Stop if there is a problem
___	II	See gaps in knowledge as interruptions to work
___	III	See gaps in knowledge as missing links in a chain
___	IV	See gaps in knowledge as missing pieces of a jigsaw
___	V	See the gaps as the most interesting part
___	VI	See gaps as pauses in the process

Set	H -	Generating a solution
___	I	Expect to be told what to do
___	II	Work out the answer from previous experience
___	III	Expect that a solution will emerge
___	IV	Resolve tasks by choosing between alternatives
___	V	Seek original solutions
___	VI	See the solution as the beginning of a new problem

Set	I -	General reactions to conflicting results
___	I	Answers should be straightforward
___	II	Options should not be discarded
___	III	Go back to the beginning if the thread is lost
___	IV	Hold a solution while developing an alternative
___	V	Discard solutions when you deem it necessary
___	VI	There are no permanent solutions

Sample Card from Card Sort Task:

Varied by color, shape, size, and number (3 x 3 x 3 x 3)

Color: red, green, blue
Shape: circle, square, triangle
Size: large, medium, small
Number: one, two, three

The rule to discover is: size and shape of symbols on placed card must match size and shape of symbols on target card.

Protocols for Administering the Career Path Appreciation Interview: Symbol Card Sort and Phrase Card Tasks

Written by Dr. T. O. Jacobs

Directions:

TODAY, WE ARE GOING TO DO TWO THINGS. FIRST, I WILL ASK YOU TO SORT SOME SYMBOL CARDS TO DISCOVER A SORTING RULE I HAVE IN MIND. NEXT, I WILL GIVE YOU SEVERAL SETS OF CARDS WITH SHORT PHRASES ON THEM. I WILL ASK YOU TO PICK THE ONES YOU ARE MOST AND LEAST COMFORTABLE WITH, AND TO TELL ME ABOUT THEM.

Symbol Card Sort Task:

NOW, I WOULD LIKE YOU TO SORT SOME SYMBOL CARDS.
[the four "target cards" are laid out]

I WOULD LIKE YOU TO SORT THESE CARDS [give the participant the stack of 162 cards, face up] ONE AT A TIME INTO FOUR PILES [point with a finger or pencil to the space just below each of the four cards, in turn]. PLEASE PLACE THE CARDS ONE ON TOP OF THE OTHER, SO ONLY THE TOP CARD SHOWS.

I HAVE A RULE IN MIND FOR HOW TO SORT THE CARDS, AND I WOULD LIKE YOU TO TRY TO DISCOVER WHAT MY RULE IS. YOUR TASK IS TO SORT 10 CARDS IN A ROW CORRECTLY.

IF YOU PLACE A CARD IN ONE OF THESE PILES [point below each of the symbol "target cards"] I WILL TELL YOU WHETHER OR NOT IT FITS MY RULE. IF YOU PLACE A CARD IN THIS PILE [point to the space below the blank "target card"], I WILL NOT GIVE YOU FEEDBACK FOR THAT CARD.

DO YOU HAVE ANY QUESTIONS?

Frequently asked questions and appropriate answers:

Can I look at the cards beneath the top card? NO.

How many cards can I use? AS MANY AS YOU LIKE.

Do I have to use them all? USE AS MANY AS YOU NEED.

Do you count the ones that go into the blank pile? YOUR TASK IS TO SORT 10 CARDS <u>IN A ROW</u> CORRECTLY.

If the participant looks uncomfortable with that answer, use the following strategy.
Ask: WHAT IS THE PURPOSE OF THE BLANK PILE?

If the participant gives the correct response, then say: YES, I COUNT THEM ALL.

If they do not know what the pile is for, then say: YOUR TASK IS TO SORT 10 CARDS IN A ROW CORRECTLY. Just repeat the instructions, and avoid giving structure to the task.

If there is no progress at about 80 cards, stop the participant and initiate a discussion of how he or she was thinking about the task.

Phrase Card Task:

I HAVE HERE SEVERAL SETS OF CARDS, EACH OF WHICH HAS A SHORT PHRASE. FOR EACH SET, I WOULD LIKE YOU TO PICK OUT THE ONE YOU ARE MOST COMFORTABLE WITH, AND THE ONE YOU ARE LEAST COMFORTABLE WITH -- AND THEN TELL ME ABOUT EACH OF THEM.

[Then hand the participant the first set of cards. It is important to let participants do the task any way they want. In particular, if they transform the task by making integrative comments, that is o.k.]

Questions to evoke discussion:

CAN YOU TELL ME ABOUT THEM?
CAN YOU TELL ME A LITTLE MORE ABOUT THEM?
CAN YOU TELL ME ABOUT THE CARDS YOU CHOSE?

In general, anything that evokes elaboration without giving structure will be acceptable. For example, avoid "explain" or "like" because these suggest a frame of reference. Extensive elaboration is not required. Generally, one or two sentences on each is enough. Most people will fall right into providing about that amount without prodding.

Appendix D

Mastermind Scoring Hierarchy

Scoring: B = BLACK PEG   W = WHITE PEG

20	B B B B	
15	B B B	
14	B B W W	
13	B W W W	
12	W W W W	
11	B B W	
10	B W W	
9	W W W	
8	B B	
7	B W	
6.5	W W	(added later because it was left out of the original hierarchy)
6	B	
5	W	

Appendix E

#______

Demographic Questionnaire

PLEASE ANSWER ALL OF THE FOLLOWING ITEMS. IF YOU DO NOT UNDERSTAND ANY OF THE QUESTIONS, PLEASE ASK THE EXPERIMENTER FOR HELP. DO NOT LEAVE ANY OF THE ITEMS BLANK.

Age __________ Gender (circle one) male female

Current graduate school major ____________________

What was your undergraduate degree(s)? ________________

How many semesters have you completed in graduate school as of December 31, 1990? ______

1. How many graduate and undergraduate quantitative courses have you had in the following topic areas? Note that quantitative courses are those that involve formulas and calculations.

   Math (e.g., calculus, algebra, geometry) __________

   Physics __________

   Chemistry __________

   Computer Programming __________

   Statistics __________

   Other:________________ __________

2. How many art courses (courses involving <u>hands-on</u> work) have you had in graduate school and undergraduate school? If you consider any of these a hobby, could you estimate hours per week you spend on average?

	# of courses	Hrs./week hobby
painting/drawing	____________	_____________

sculpting/pottery	________	________
music (instrumental)	________	________
music composition	________	________
architecture	________	________
other: ________	________	________

3. In the past, were you seriously involved in any of the activities listed above? For example, you may have played an instrument for many years in your childhood, but you don't currently play as a hobby or take courses. Please describe.

____

____

____

Note: The demographic questionnaire appeared to participants as one page with smaller margins and single spacing, instead of the two pages shown here.

Appendix F

Test Anxiety Profile

This is a test about your thoughts and feelings toward taking tests. There are no right or wrong answers. We are interested in what your thoughts and feelings are as you go through the testing situation today. Consider your current feelings when answering questions, not how you have felt in the past or how you anticipate you will feel.

SAMPLE ITEM

You might see an item such as the following one. The word in parentheses (Me) identifies what you should describe. Mark only one X on the line that best describes your feelings.

HOW I FEEL WHILE TRYING TO TAKE A TEST UNDER NOISY CONDITIONS

(Me)

TENSE ____:____:____:____:____:____:____ RELAXED

If you feel more "relaxed" than "tense" when you take a test under noisy conditions, you would have marked an X closer to the "relaxed" side of the scale, such as:

TENSE ____:____:____:____:____:__X_:____ RELAXED

Please place your marks on the lines, not between them.

If you have no questions, please proceed to the next page.

HOW I FEEL WHILE TAKING THIS TEST RIGHT NOW

(Me)

CALM ____:____:____:____:____:____:____ JITTERY

(Fingers)

STIFF ____:____:____:____:____:____:____ RELAXED

(Me)

HELPLESS ____:____:____:____:____:____:____ SECURE

(Breathing)

LOOSE ____:____:____:____:____:____:____ TIGHT

(Me)

WORRIED ____:____:____:____:____:____:____ CAREFREE

WHAT MY THOUGHTS ARE LIKE RIGHT NOW

(Ideas)

CLEAR ____:____:____:____:____:____:____ CONFUSED

(Me)

UNSURE ____:____:____:____:____:____:____ SURE

(Prepared)

UNREADY ____:____:____:____:____:____:____ READY

(Thoughts)

JUMBLED ____:____:____:____:____:____:____ EASY

(Mind)

WORKING ____:____:____:____:____:____:____ BLANK

Appendix G

Frequency Distributions of Variables

An examination of the frequency data plotted on a bar graph show for all CPA measures, the distributions tended to follow a normal curve (Figures 8-10).

For the creativity measure, KAI, the median score of 104 is above the norm of 96 for the measure (Figure 11). This indicates that this study's sample tended to be more "innovator" than "adaptor."

The Mastermind performance (Figure 12) had a mean score of 21, a median of 23.3, and a range from 7 to 27.

The GRE Verbal distribution (Figure 13) followed a normal curve, with the addition of an outlier at the high end. The range was from 340 to 800, with a mean of 518. The GRE Quantitative distribution of scores (Figure 14) seemed to be positively skewed with a mean of 521 and a range of 350 to 800. The GRE Analytical scores (Figure 15) were distributed somewhat normally, with a range of 350 to 800 and a mean of 552.

The measures of anxiety, FA and TI, tended to follow a symmetric unimodal distribution (see Figures 16 and 17).

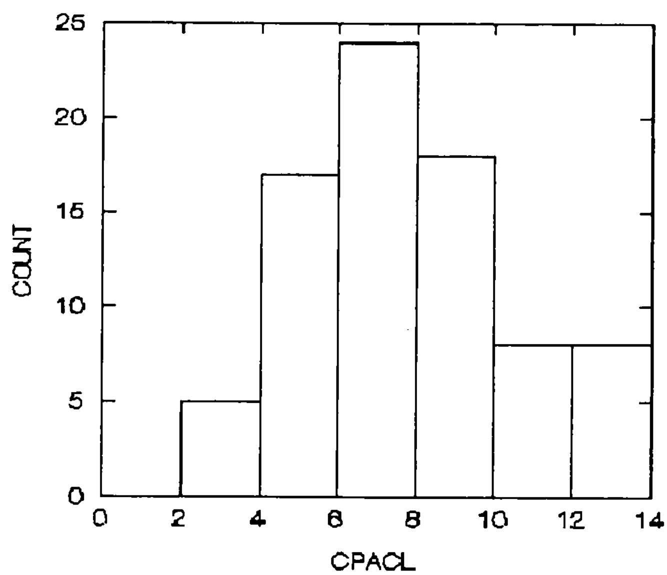


Figure 8. Career Path Appreciation: A frequency distribution of participants' current level scores.

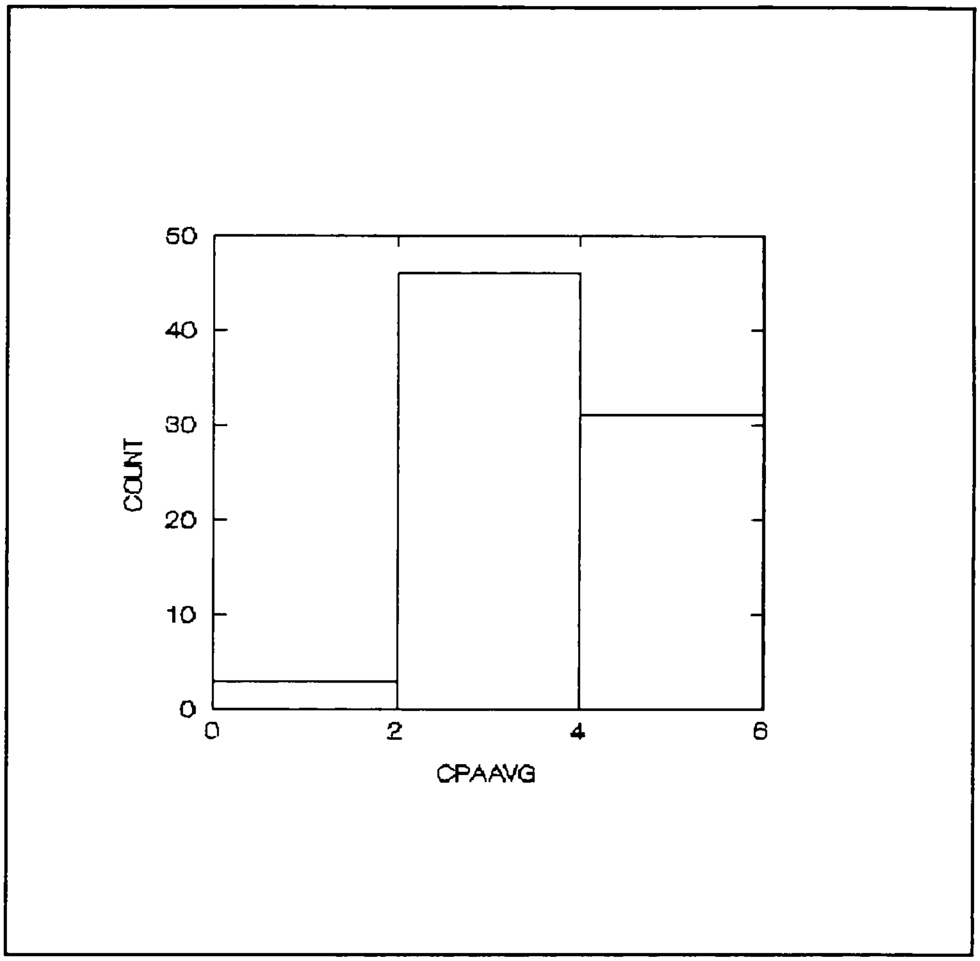


Figure 9. Career Path Appreciation: A frequency distribution of participants' phrase card average scores.

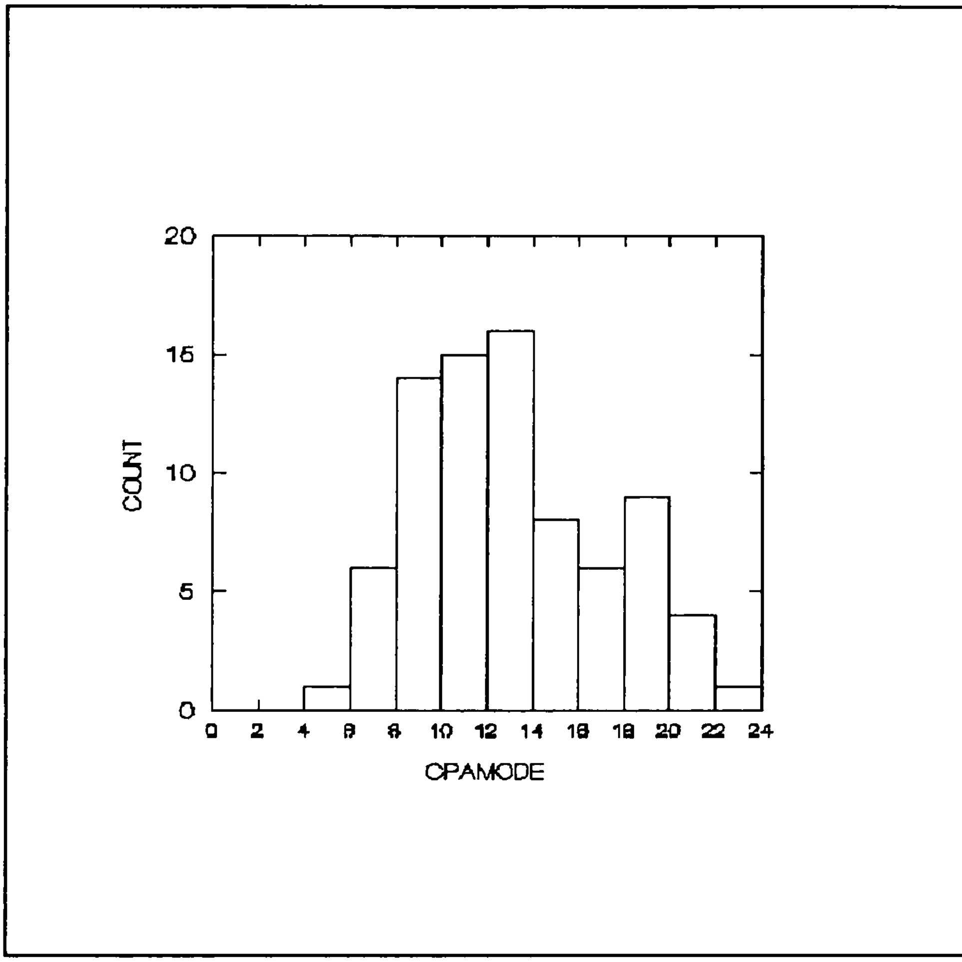
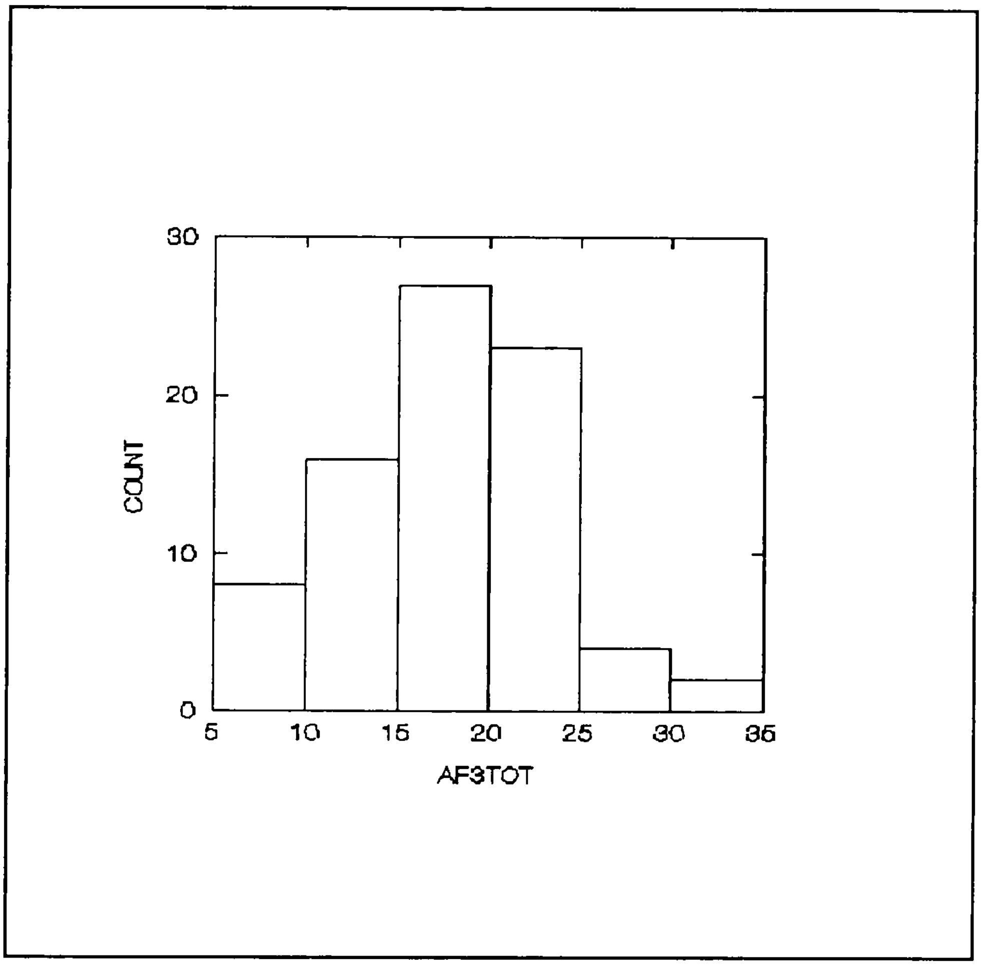


Figure 10. Career Path Appreciation: A frequency distribution of participants' mode scores.



<u>Figure 11</u>. Test Anxiety Profile: A frequency distribution of participants' feelings of anxiety (FA) scores.

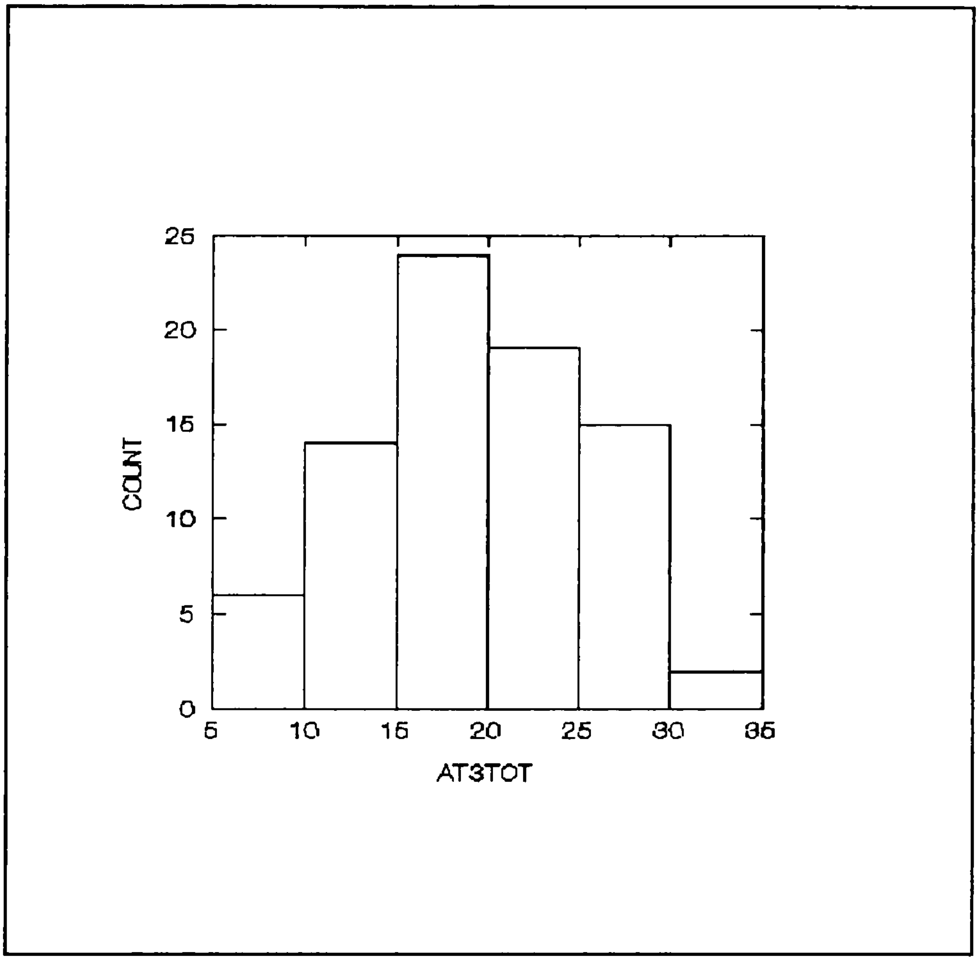


Figure 12. Test Anxiety Profile: A frequency distribution of participants' thought interference (TI) scores.

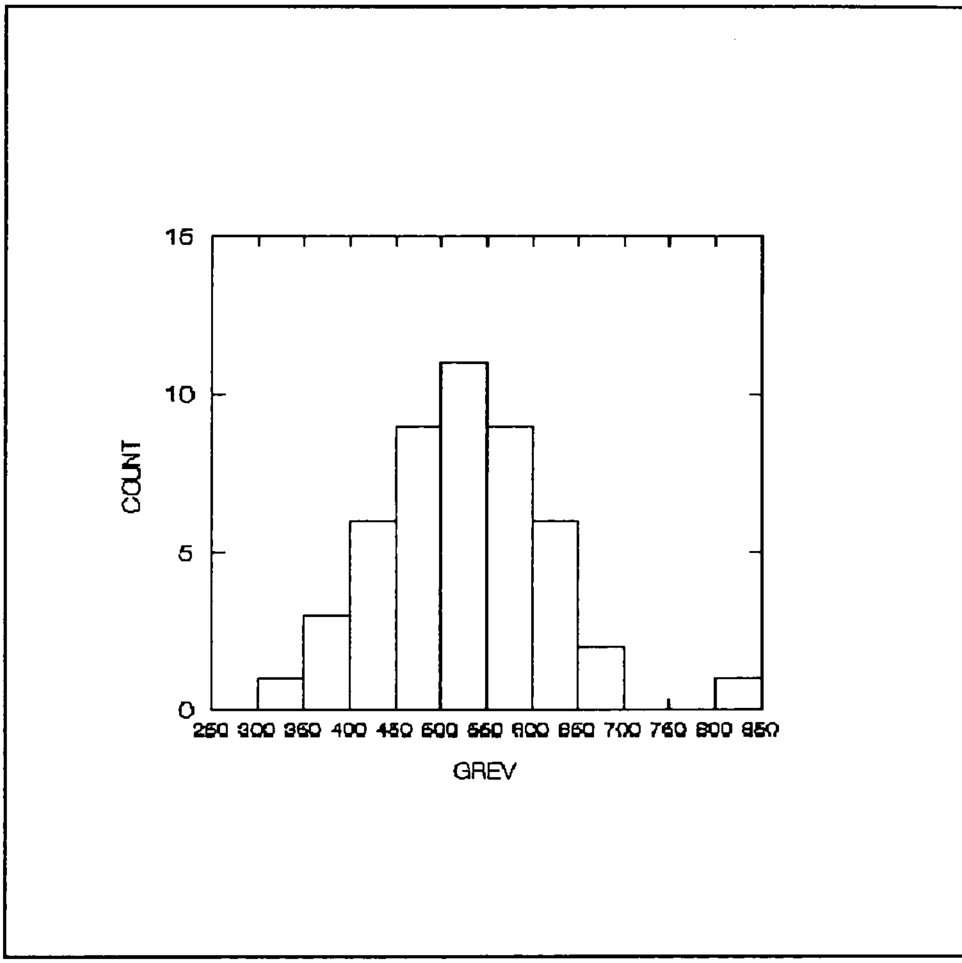
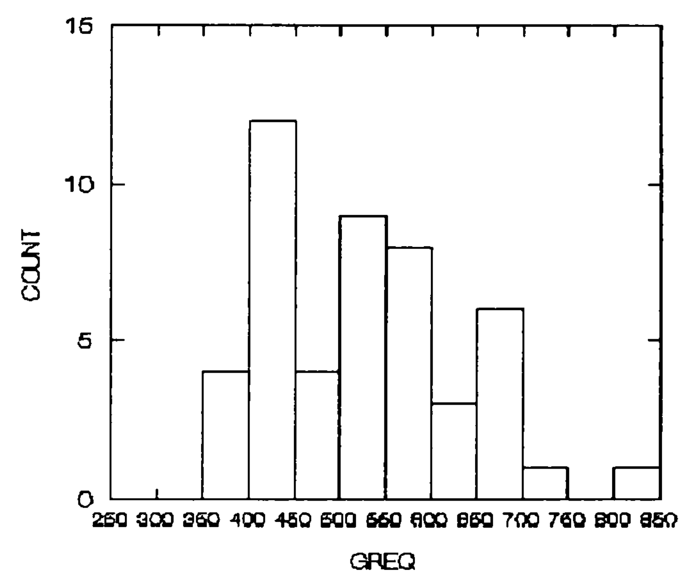


Figure 13. Graduate Record Examination: A frequency distribution of participants' Verbal scores.



<u>Figure 14</u>. Graduate Record Examination: A frequency distribution of participants' Quantitative scores.

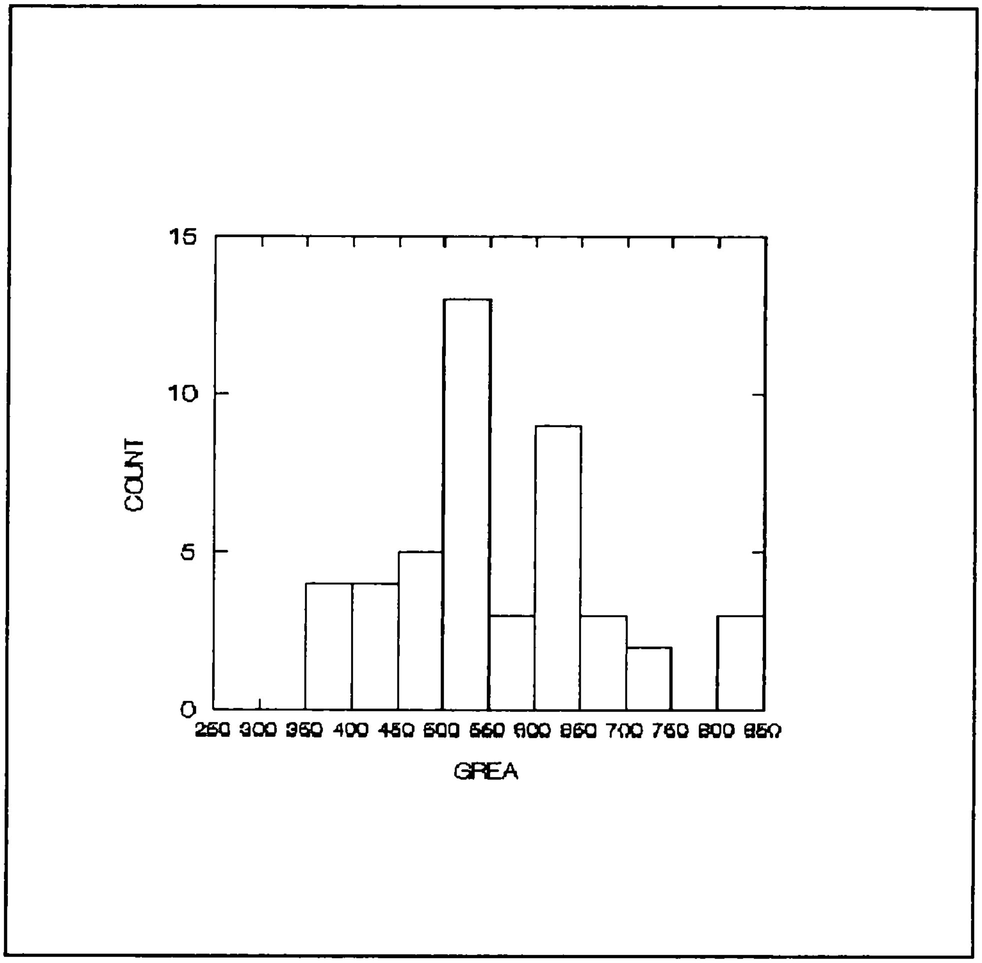


Figure 15. Graduate Record Examination: A frequency distribution of participants' Analytical scores.

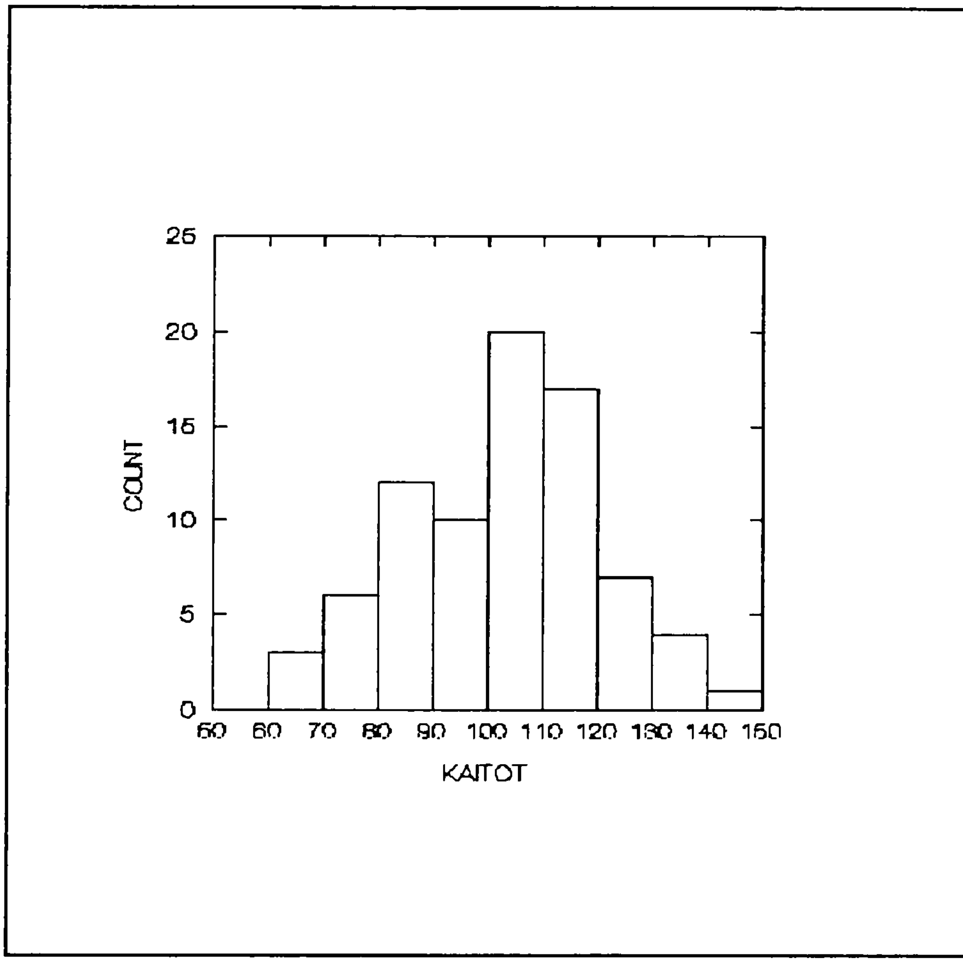


Figure 16. Kirton Adaptation-Innovation Inventory: A frequency distribution of participants' scores.

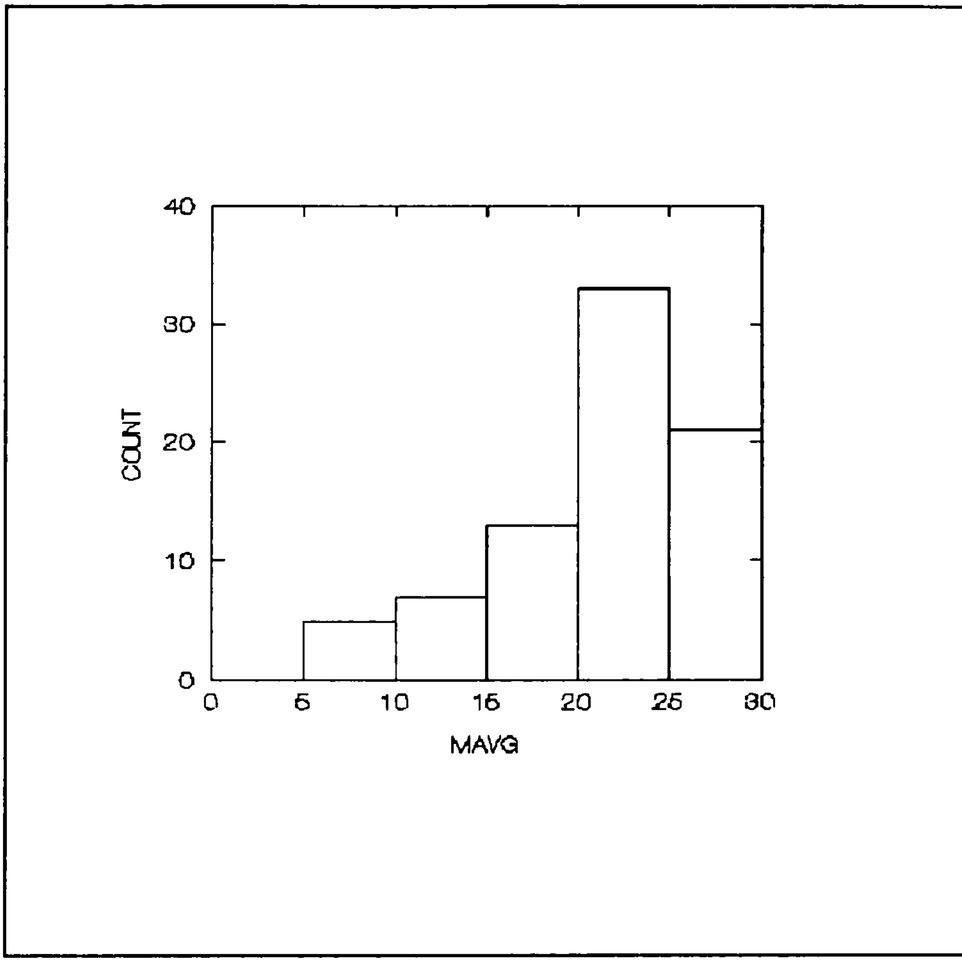


Figure 17. Mastermind: A frequency distribution of participants' scores.

Appendix H

Additional Results Not Previously Hypothesized

Additional findings are presented here. In some cases, these findings show interesting pathways to future research, and in others, support is given for the present investigation.

The r_s for Level 2 and the r_s for Level 3 were significantly correlated (r = .318, p < .005). This suggests that there is a tendency for a participant to be similarly linear in Levels 2 and 3. The r_a and r_m for Level 3 were significantly correlated (r = .943, p < .000) indicating a higher achievement for participants with better matching strategies. The r_a and r_s for Level 3 were also significantly correlated (r = .351, p < .002) indicating higher achievement for participants with linear strategies (in a nonlinear task!)

The CPA current level of a person was positively correlated with age (r = .458, p < .000). The CPA mode was positively correlated with the Analytical GRE score (r = .334, p < .023) and the number of hours per week spent at art (r = .266, p < .017). And both CPA scores (current level and mode) were positively correlated with the KAI (current level r = .299, p < .007; mode r = .297, p < .007).

The GRE scores were significantly correlated with several measures, but due to missing GRE scores (scores for 32 of the 80 participants were unobtainable), the findings should be considered carefully. The Mastermind score positively correlated with all three GRE scores (GRE Verbal r = .420, p < .003; GRE Quantitative r = .644, p < .000; GRE Analytical r = .511, p < .000). The number of hours spent per week on art were negatively correlated with the Verbal and Analytical scores (GRE Verbal $r = -.291$, $p < .045$; GRE Analytical $r = -.289$, $p < .052$). The total number of quantitative classes reported (from undergraduate and graduate degrees) correlated positively with the Quantitative GRE score ($r = .304$, $p < .036$), and gender correlated positively with the Quantitative GRE scores ($r = .322$, $p < .021$). Gender correlated negatively with thought interference ($r = -.241$, $p < .032$) indicating that women recorded more anxious thoughts about the task at Time 3. This information was used in order to understand the finding that men had significantly higher achievement and matching indices for Level 2 than women. A partial correlation was calculated, as reported in Chapter 3, which partials out the effects of the significant correlations among Quantitative GRE scores, thought interference anxiety, and gender.