METACOGNITION AND DECISION MAKING: A BRUNSWIKIAN ANALYSIS

by

PETER SAMUEL SCHAEFER

(Under the Direction of Robert Mahan)

ABSTRACT

Cognitive Continuum Theory (Hammond et al., 1987) posits that differing task characteristics in multiple cue probability learning environments can induce either analytical or intuitive cognition. The different modes of cognition are in turn associated with differing patterns of performance, including the ability to declare how one combines information when making judgments (termed *insight*). Insight is thus largely a state phenomenon, or one that is constrained or fostered by task characteristics. Conversely, metacognition (one aspect of which involves the ability to declare how one has enacted a strategy) is often thought of as an individual difference. Eight different environments in two experiments were constructed to induce either analytical or intuitive cognition. The environments were varied along several different dimensions, including whether the relationship between cues and criterion values was linear or nonlinear. In addition to completing a judgment task in one of these eight conditions, all participants completed two standard measures of metacognition. In this manner, the contributions of both task characteristics and individual differences to performance were assessed. Results indicated that both task characteristics and individual differences were related to performance.

INDEX WORDS: Cognitive Continuum Theory, Metacognition, Decision Making, Multiple Cue Probability Learning

METACOGNITION AND DECISION MAKING: A BRUNSWIKIAN ANALYSIS

by

PETER SAMUEL SCHAEFER

B.S., Valdosta State University, 1997

M.S., The University of Georgia, 2000

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2003

METACOGNITION AND DECISION MAKING: A BRUNSWIKIAN ANALYSIS

by

PETER SAMUEL SCHAEFER

Major Professor: Robert Mahan

Committee: Zachary Estes
Adam Goodie

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2003

ACKNOWLEDGEMENTS

I would like to extend my gratitude to those who have helped me in this long, and at times seemingly interminable, journey. First, I would like to thank my major professor, Robert Mahan, for allowing me to work in his lab on a series of fascinating projects. Secondly, I would like to thank the other members of my doctoral committee: Zachary Estes, for his thorough criticism and valuable time, and Adam Goodie for his constant encouragement. I would also like to thank some of the graduate students who have given me invaluable feedback and support, along with much laughter: Cristina Williams, for the intellectual companionship and wonderful discussions; Lawrence Michels, for a refreshing and affirming view of the world; and perhaps most of all, Jennifer “I didn’t even know her” Solberg.   While I continue to doubt your sanity, I don’t doubt the value of having known you. To my mother, whose unflagging support was indispensable. To my brother, for setting an example to follow. Finally, I should mention someone in whose intellectual shadow I have often labored: my father, German Diego Schaefer.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Terminology

Throughout this paper, the initial usage of each term is italicized. Terms are discussed in the main body of the paper in nontechnical language. More formal definitions of all italicized terms are given in Appendix A. All definitions are paraphrased from Cooksey (1996).

Probabilistic Functionalism

One of the basic goals of research in judgment and decision making is that of specifying general rules which describe the decision making process. However, much of the early psychological research on judgment has been criticized for focusing on the environment at the expense of the organism (Doherty, 2001). The concept of *probabilistic functionalism* (Hammond, 1966; Brunswik, 1956a) epitomizes this concern with the interaction between an organism and its environment.

Probabilistic functionalism focuses on decision making in environments that are more typical of everyday ecologies. For example, judgments are often based on information that is less than perfectly predictive of some event of interest. In such situations, one earmark of adaptive behavior is the ability to base judgments upon information to the extent that the information is predictive (Brunswik, 1943). Further, judgments are often based upon multiple pieces of information. (An extended discussion of probabilistic functionalism is given in Appendix B.)

Multiple Cue Probability Learning

Out of the concept of Probabilistic Functionalism arose the paradigm of Multiple Cue Probability Learning (MCPL). In a typical MCPL task, an organism is given several pieces of information (known as *cues*). The organism learns, in a trial-by-trial fashion, to utilize the cues to estimate some characteristic of the environment (known as the *criterion*). There are two kinds of feedback typically used in an MCPL experiment: *outcome feedback*, and *cognitive feedback*.

Outcome feedback is given on a per-trial basis, and consists of nothing more than supplying the correct criterion value that accompanies some set of values, following an estimation of that criterion value by the organism. Cognitive feedback, in contrast, occurs at the end of a block of training trials. Cognitive feedback is sometimes numerical in nature, such as providing the means and standard deviations of an organism's judgments across a block of training trials, accompanied by the mean and standard deviation of the correct criterion values across that block of trials.

More effective, however, are graphs which illustrate aspects of the ecology as well as the manner in which the cues were utilized by the organism. For example, it is common to show graphs depicting the relationships between cue values and criterion values. It is also common to show graphs depicting the relationships between cue values and judgments made (Balzer, Doherty, & O' Connor, 1989; Doherty & Balzer, 1988). (More detailed findings from the MCPL literature are given in Appendix C.)

The Lens Model Equation

The Brunswikian concern with the necessity of a symmetrical emphasis evolved into the Lens Model Equation (LME). The LME is depicted graphically in Figure 1. The mathematical underpinnings of the LME revolve around ordinary least squares regression. The symmetrical emphasis placed upon the organism and the ecology can be seen in the analogous regressions applied to both (Hammond, 1955). The choice of regression and correlation was deliberate, because both statistical concepts involve inevitable uncertainty. In this manner the probabilistic nature of decision making is incorporated in the LME.

The left hand side of the LME figure represents the ecology. There were a total of eight ecologies across the two experiments, corresponding to the eight experimental conditions. When the observed criterion values were regressed upon the cue values (generating the *ecological regression model*), a set of predicted criterion values was generated. The multiple correlation between the observed criterion values and the cue values expressed the *predictability* in an environment. The higher the predictability of an environment, therefore, the smaller the discrepancy between the observed and predicted criterion values.

The right side of the column represents the judgment processes of an organism. Regressing the judgments made by the organism upon the cues yielded the organism's *policy*. The computation of a policy allowed for the generation of a set of predicted judgments, in the same fashion that computation of the ecological regression model allowed for the generation of a set of predicted criterion values.

Out of this process, three LME performance indices were created. The first was that of *achievement*. Achievement is the simple Pearson correlation between the judgments made by an organism, and the observed criterion values. It is thus a measure of empirical accuracy. The second performance index created was that of *cognitive control* (Hammond & Summers, 1972). This is the correlation between the judgments made by an individual and the predicted judgments generated resulting from the individual's policy. Higher levels of cognitive control indicate greater consistency in the application of a policy. The final performance index was that of *matching*, which is the correlation between the predicted criterion values and the predicted judgments of an organism. Higher levels of matching indicate greater correspondence between the correct rule for a task and the policy enacted by an individual. (For a more detailed explication of the Lens Model Equation, see Appendix D.)

Cognitive Continuum Theory

Philosophers and psychologists have long held the view that analysis and intuition were two disparate modes of cognition (Kahneman & Tversky, 1982). However, the manner in which analysis and intuition have been defined is problematic. While analysis has been fairly well described, intuition has often been described as something that was not analysis (Brooks, 1978; Beach & Mitchell, 1978).

In contrast, Cognitive Continuum Theory (CCT) builds upon Brunswik's thesis that analysis and intuition are not disparate modes of cognition, but rather opposing poles of a continuum (Brunswik, 1956b). CCT contributes to the research in this area by, firstly, specifying the types of tasks that should give rise to analytical or intuitive cognition. Secondly, CCT defines the differing patterns of performance that should accompany analytical or intuitive cognition.

The characteristics that differentiate between intuitive and analytical tasks are listed in Table 1. These characteristics can be used to assign a condition a Task Continuum Index (TCI) score. Higher TCI scores indicate a more analytical task. In this fashion, CCT allows for the a priori classification of tasks as more analytical or more intuitive.

Of equal importance are the differences in performance that arise due to engagement in either analysis or intuition. These differences are shown in Table 2. The difference involving *insight*, or degree of correspondence between an organism's stated policy and statistically captured policy, is of major theoretical importance, and its conceptual link to metacognition is explored in the next section in detail. Insight has been measured in various ways. However, the most commonly used method for measuring insight was inappropriate in the current study. Therefore, another method of measuring insight was used (see Appendix E for details).

Table 1. Task Characteristics in Cognitive Continuum Theory

Task Characteristics	Intuition Inducing	Analysis Inducing
Number of cues	Large (greater than 5)	Small (2-4)
Cue intercorrelation	High	Low
Degree of predictability	Low	High
Cue-Criterion Relationships (a.k.a. *function forms*)	Linear	Nonlinear
Variation among beta scores	Smaller	Larger

Table 2. Cognitive Mode Differences in Cognitive Continuum Theory

Intuitive Cognition	Analytical Cognition
Low Insight	High Insight
High confidence in performance	Low confidence in performance
Low confidence in insight	High confidence in insight
Rapid information processing	Slow information processing
Errors are normally distributed	Errors are few and large
Low cognitive control	High cognitive control
High achievement unlikely	High achievement likely

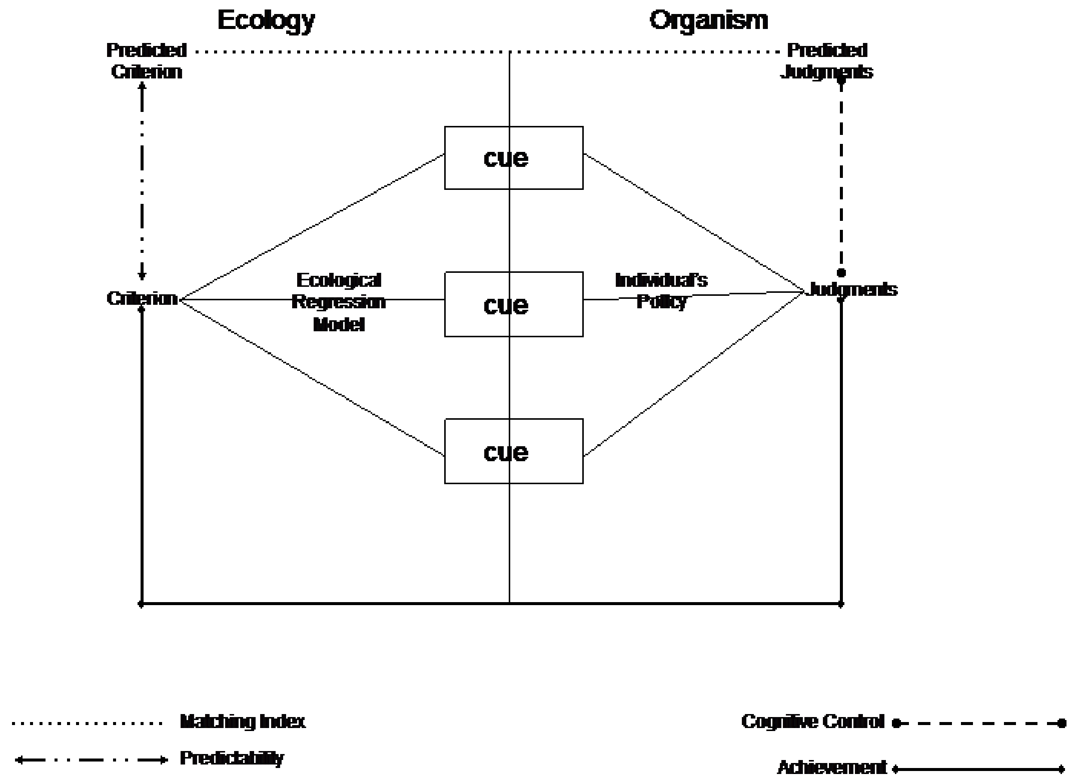

Figure 1. Lens Model Equation

CHAPTER 2

METACOGNITION AND INSIGHT

Metacognition has been defined in various ways by various researchers. It has been defined as "…awareness of how one learns, when one does and does not understand, and the assessment of progress both during and after performance" (Gourgey, 1998), and "….an individual's conscious awareness and control of the cognitive processes involved in attending to and focusing on processing, comprehending, and remembering information" (Schmitt & Baumann, 1986). One early definition proposed by Flavell (1979), while the broadest, is perhaps the best: he defined metacognition as "knowledge and cognition about cognitive phenomena; metacognitive knowledge consists primarily of knowledge or beliefs about what factors or variables act and interact in what ways to affect the course and outcome of cognitive enterprises".

Central to most of the definitions of metacognition is the conception of metacognition as an ability, greater amounts of which allow one to more accurately assess performance. This assessment may be of a predictive nature, such as spending greater amounts of time studying materials that are not well understood (Bisanz, Vesonder, & Voss, 1978), or of a post-hoc nature, such as the degree of confidence one expresses in the ability to subsequently recall learned material (Nelson & Dunlosky, 1991; Flavell, Freidrichs, & Hoyt, 1970; Maki, Foley, Kaijer, & Thompson, 1990).

Sometimes self-ratings of the extent to which individuals engage in metacognitive behaviors (i.e., comprehension checking during performance, assessment of strategy success) are utilized. One such measure is the Metacognitive Awareness Inventory (Schraw, 1994). In other cases, the extent to which individuals can display declarative knowledge is used as an indication of metacognition (Coleman & Shore, 1991; Swanson, 1990; Artzt & Armour-Thomas, 1996). Declarative knowledge involves being able to state what strategy one is using and to articulate the steps to others.

Perhaps the most commonly used method to assess metacognition, however, is the ability to reflect upon an answer that has been chosen and assign a level of confidence in the correctness of that answer (Schraw & Roedel, 1994; Schaefer, Williams, Goodie, & Campbell, 2002). Higher levels of metacognition are indicated by smaller discrepancies between average confidence and average performance (known as *accuracy scores*, which are difference scores reflecting absolute magnitude of difference; Morris, 1990). [1]

Metacognition in MCPL Tasks

There has been at least one explicit link made between the metacognition and MCPL literatures. A well-established finding in the metacognition literature is the "hard-easy" effect (Lichtenstein & Fischoff, 1977). Greater levels of metacognition are seen when individuals are said to be well calibrated; that is, there is little or no difference between average confidence and average performance (Suantak, Bolger, & Ferrell, 1996).

However, factors that affect performance more than confidence will result in miscalibration. It has been found that, as test items become more difficult, performance drops and overconfidence results; that is, confidence exceeds performance (Schraw & Roedel, 1994). Conversely, as test items become normatively easier, performance

[1] This means that accuracy scores should be negatively correlated with performance. However, for ease of interpretation, the valence of all accuracy score correlations was reversed.

increases and underconfidence results (Schraw, Potenza, and Nebelsick-Gullet, 1993).

The normative difficulty of the test items imposes a ceiling effect upon performance. Within an MCPL task, an analogous manipulation can be accomplished by raising or lowering the predictability of the task ecology. If predictability (the amount of variance in the criterion explained by cue variance) is raised, the environment becomes easier in that the appropriate rule, once learned, explains more variance. Conversely, when predictability is lowered, the environment becomes more difficult. Not only is it harder for an individual in such a situation to extract the appropriate rule for combining and utilizing cue information, but the extent to which such utilization will yield correct answers has been minimized.

Just such an experiment was conducted by Doherty, Brake, and Kleiter (2001). Individuals in an MCPL experiment were asked to make a series of judgments. Following each judgment, individuals were asked to indicate how confident they were that they had chosen correctly. When predictability in the task ecology was lowered, the "hard" effect was seen. Namely, overconfidence increased. When predictability was increased, the "easy" effect appeared: underconfidence was seen.

The examination of metacognition and performance within the MCPL paradigm appears to offer several advantages over traditional methods of examining the relationships between metacognition and performance. For example, although some authors (Koriat, 1997) have speculated about the role that cues play in metacognition, cues in traditional metacognition research are thought to be largely idiosyncratic (e.g., subjective familiarity with a cue; Metcalfe, Schwartz, & Joaquim, 1993) and therefore not easily amenable to manipulation and description.

However, in an MCPL task the extent to which individuals utilize cue information (i.e., weighing cues that are weighed more in the task ecology) can be directly examined, as the cues are made explicit. In addition, MCPL tasks carried out within the Brunswikian tradition allow for the derivation of several indices of performance, such as the extent to which answers correspond with criterion values, the extent to which the rule utilized by an individual corresponds to the correct one for that ecology, and how consistently a strategy is applied by an individual.

Finally, as been noted elsewhere, metacognition and Brunswikian LME research are a good fit: "Calibration research is, in a fundamental way, akin to research based on Brunswik's Lens Model. In both lines of research, the focus is on empirical accuracy, or correspondence between judgments and environmental outcomes" (Doherty, Brake, & Kleiter, 2001, p. 317). For Brunswik, an earmark of adaptive behavior was the utilization of cues to the extent to which they were predictive of some event (Brunswik, 1943). The consideration of metacognition as a potentially relevant individual difference variable underscores the attention paid to the organism as well as the environment in LME research.

Metacognition and Insight

Although insight is an important aspect of Cognitive Continuum Theory, at times the argument for its existence has been circular. It seems that it has often been argued as having been demonstrated because cognitive control was high, and that because cognitive control was high, so was insight (Hammond & Summers, 1972).

Furthermore, when insight has been measured, it has usually been done outside the context of CCT research. More common is the examination of insight in tasks that would be considered primarily intuitive (e.g., positive linear function forms, less-than-perfectly predictable environments; Reilly & Doherty, 1992; Brehmer, 1977; Reilly & Doherty, 1989; Reilly, 1996) with little systematic manipulation of task factors thought to induce the modes of cognition at the poles. Rather, one variable was manipulated and performance between conditions compared (e.g., predictability; Doherty, Brake, & Kleiter, 2001; Gray, 1979).

One recent exception is that of Haarbauer (1996). Unfortunately, the main prediction of CCT that insight would be higher for persons in an analytical task was not borne out, although insight was significantly correlated with performance in both the analytical and intuitive tasks. Haarbauer suggested that part of the problem might be that many studies of this sort are within-subjects in nature, and that a between-subjects design might not achieve the same results unless possibly relevant individual difference information were also collected. In addition, Haarbauer did not manipulate the degree to which the cue-criterion relationships (a.k.a. *function forms*) were linear or nonlinear.

There are grounds for positing that insight and metacognition are similar constructs, and hence metacognition may be put forward as one such potentially relevant individual difference. Insight is essentially a measure of the ability to describe the process by which one combined information to make a set of judgments. In metacognition, one of the most prominent concepts is declarative knowledge, or the ability to verbalize the process by which one is completing some task (e.g., solving physics problems; Coleman & Shore, 1991).

Insight into the process by which one integrates cue information and makes judgments is thought to be related to the ability to benefit from feedback (Hammond, 1996). That is, the ability to reflect in some fashion upon both what one has been doing and what one should be doing allows one to minimize the discrepancy between the two behaviors. Insight is also theoretically related to the ability to consistently apply a policy (i.e., exhibit high cognitive control; Hammond, 1990). Similarly, a major assumption of metacognitive theory is that declarative knowledge also underlies the ability to both consistently apply a strategy (Manning & Payne, 1996) and to benefit from feedback (Pressley & Ghatala, 1990; Butler & Winne, 1996).

Finally, the research literature indicates that the relationship between predictability and metacognition may parallel that posited by CCT to exist between predictability and insight. For example, Doherty and Kleiter (2001) found that, as a task was made less predictable, metacognition became less accurate. Similarly, CCT predicts that insight should drop as predictability is lowered, because the task becomes more intuitive as predictability is minimized.

CHAPTER 3

EXPERIMENT ONE

Purposes and Rationale

The purposes of the first experiment were to examine the relationships between performance in MCPL tasks, cue-criterion function form, insight, and metacognition.

CCT and Function Forms

Although the manipulation of function forms is theoretically important for CCT, like insight it has not often been examined. The more common approach has been to utilize linear function forms for both analytical and intuitive tasks while manipulating the other task characteristics that are thought to induce either analytical or intuitive cognition. Further, in the one experiment that examined nonlinear function forms (Hammond, Hamm, Grassia, & Pearson, 1987), the nonlinear function form was merely described as "a stochastic function" with no further details as to the nature of the function.

Function forms are theoretically important because CCT posits that linear cue usage (i.e., intuitive cognition) is the default mode for humans. Linear models describe many activities in the animal kingdom quite well (see Appendix F for a more detailed discussion of linear models of judgment). These models are almost invariably, however, concerned with the processing of positive linear relationships. For example, it has been found that positive linear relations are much easier to learn than negative linear ones (Bjorkmann, 1965).

Further, there is also reason to assume that the type of nonlinear relationship affects performance. As noted in the discussion of MCPL research (see Appendix C), there appears to be a definite hierarchy of complexity when it comes to cue-criterion relationships. Although CCT predicts that nonlinear cue-criterion relationships should result in analytical cognition and, hence, higher levels of performance, MCPL research indicates that individuals have an easier time learning linear cue-criterion relationships.

More specifically, it appears that individuals find the cue-criterion function forms, from easiest to most difficult, to be: positive linear, negative linear, U-shaped quadratic, inverted U-shaped quadratic (see Appendix C). Hammond (1996) has argued that the “automatic” nature of linear processing in humans can be contrasted to the “controlled” process necessary when learning nonlinear function forms. If this is correct, then it made sense to utilize the inverted-U shaped cue-criterion relationship for the nonlinear tasks, as well as the positive linear cue-criterion relationship for the linear tasks. In essence, it was expected that the largest difference in performance would occur between these two function forms.

CCT, Metacognition, and Insight

It has been argued that metacognition and insight are similar constructs in that both are thought to be correlated with performance, and both are measures of declarative knowledge. However, the manner in which metacognition was measured in the current study is dissimilar in that of the aforementioned Doherty and Kleiter (2001) study. There were several reasons for this dissimilarity.

Firstly, measuring metacognition *within* an MCPL task was problematic for the current study, which was attempting to draw a comparison between the constructs of insight and metacognition. This was because inserting questions into a task has been shown to *increase* metacognitive activity (Stimson, 1998). Therefore, measuring metacognition in such a manner might artificially inflate metacognition and/or insight.

Secondly, the use of standard measures of metacognition allowed for the placement of the findings within the metacognition literature. It is normally the case that scores on standard metacognition measures are elicited from participants, and the relationships between those scores and various measures of performance (e.g., GPA) are examined. That is, the measurement of metacognition is usually not task specific.

Thirdly, the use of standard measures also allowed for some current concerns within metacognition research to be addressed. For example, a recent goal has been the examination of the extent to which metacognition, as measured on such standard measures, is either a domain-general or a domain-specific phenomenon (Schraw, Dunkle, Bendixen, & Roedel, 1995; Schraw, 1997; Schraw & Nietfield, 1998). The manner in which metacognition was measured in the present study attempted to extend the examination of metacognition to a type of task not examined in such a way before.

Further, investigations concerning the domain-generality or domain-specificity of metacognition have been largely concerned with more "academic" domains such as history (Schraw & Nietfield, 1998), or knowledge of computer programming (Veenman, Elshout, & Meijer, 1997). Conversely, many judgment researchers claim that MCPL tasks are much more akin to tasks in "everday life" than many other tasks used to study decision making (Hammond, 2001). If this claim is correct, then the examination of how metacognition is correlated with performance in such domains would be an important generalization.

Fourthly, the measurement of metacognition in such a manner allowed for the independent assessment of the contributions to performance accrued from insight and metacognition. It also allowed for an assessment of the relationship between the two constructs of metacognition and insight. The difference between how the two constructs were measured can be couched in terms of trait versus state.

Essentially, metacognition as it was measured was as a trait (within-subjects) construct, one that was independent of the task characteristics. This was accomplished by assessing metacognition via standard individual difference measures before the judgment task began. Conversely, insight within the context of CCT is unavoidably a state (between-subjects) phenomenon. That is, CCT specifically predicts that task characteristics will affect the level of insight exhibited by an individual. Finally, the manner in which metacognition and insight were measured allowed for a more stringent test of the proposed relationships among the measures of performance, insight and metacognition. The manner in which these proposed relationships were tested is described in the next section.

Hypotheses

In accordance with CCT, insight was predicted to differ from condition to condition. The more analytical the task, the higher the level of insight should be. Based upon the characteristics listed in Table 2, Task Continuum Index (TCI) scores were computed for all four tasks. The TCI scores for the four tasks are given in Table 3. A higher TCI score indicated a more analytical task. The TCI scores were used to guide all hypotheses.

Similar usage of TCI scores has been able to predict quite accurately at least the order of effects (Hammond, Hamm, Grassia, & Pearson, 1987). However, the aforementioned article did not test the null hypothesis but merely the order of effects. Therefore, it was possible that while significant differences may not result, the order of correlational magnitude would be in the direction proposed (i.e., there will be a significant positive rank-order correlation between performance and the TCI scores).

This reflects an underlying limitation of the TCI, which claims to be ordinal but not interval in nature; that is, just because two tasks with TCI scores of 2 and 4 might display significant differences in performance, it does not necessarily follow that tasks with TCI scores of 4 and 6 would also do so. Therefore, supplemental analyses included Spearman-rho rank-order correlations between TCI scores and the dependent variables.

Further, the TCI contained no means by which to weight differentially the factors that contribute to analytical cognition induction. That is, the factors of beta weight variability, average cue intercorrelations, number of cues, and cue-criterion function forms are equally weighted. Hammond, Hamm, Grassia, and Pearson (1987) have stated that this was done to underscore the compensatory nature of Brunswikian decision making (see Appendix D for a discussion of the LME as a compensatory model).

However, it was plausibly the case that function forms would be more important in the induction of analytical cognition than indicated by the TCI. This was especially so because, as mentioned earlier, CCT research has seldom manipulated the nonlinearity of function forms. Conversely, MCPL research has shown that function forms constitute an important variable.

The TCI scores, shown in Table 3, guided the hypotheses. In each hypothesis, the order of means could be restated to say that a significant main effect was predicted in each case for the task type, in which analytical tasks are expected to result in higher performance than intuitive tasks (with TCI means of 9.22 versus 6.46, respectively). Similarly, quadratic tasks should result in higher performance than linear tasks (with TCI means of 11.38 versus 2.15, respectively). Finally, for each hypothesis, there was no predicted interaction effect. There were 11 hypotheses tested in the current experiment.

The influence of insight and metacognition upon the Lens Model Equation performance indices of matching, accuracy, and cognitive control were explored utilizing correlations and, when appropriate, simultaneous regression equations. Because the operating assumption was that performance would be affected by insight and metacognition, the MANOVA results were used as a "gateway" analysis to indicate which conditions would be so examined. For example, if a significant effect was seen for function form, the differing correlational patterns between performance, insight, and metacognition in the linear and nonlinear tasks would be examined.

Hypothesis 1	It was predicted that performance (achievement, matching, & cognitive control) would be, from highest to lowest, in the following order: analytical-quadratic (A-Q), intuitive-quadratic (I-Q) analytical-linear (A-L), and intuitive-linear (I-L)
Hypothesis 2	It was predicted that insight would be, from highest to lowest, in the following order: A-Q, I-Q, A-L, I-L.
Hypothesis 3	It was predicted that the difference scores between insight and performance (achievement, matching, & cognitive control) would be, from smallest to largest, in the following order: A-Q, I-Q, A-L, I-L
Hypothesis 4	It was predicted that the difference scores

	between metacognition and performance (achievement, matching, cognitive control) would be, from smallest to largest, in the following order: A-Q, I-Q, A-L, I-L
Hypothesis 5	It was predicted that the difference scores between metacognition and insight would be, from smallest to largest, in the following order: A-Q, I-Q, A-L, I-L

H1: CCT predicted that as task properties became more and more likely to induce analytical cognition, performance indices would increase. Since TCI scores were computed to indicate the extent to which task properties were more (or less) likely to induce analytical cognition, the rank order of the TCI scores and the performance indices should be the same. Thus, it was predicted that the Spearman-rho rank-order correlations between the TCI scores and the performance indices would be significant and positive. MANOVA results were predicted to indicate that performance was higher in the analytical than intuitive conditions, and higher in the quadratic than linear conditions.

H2: CCT predicted that insight would be fostered as tasks became more analytical and suppressed as tasks became more intuitive. The rank order of the TCI scores and insight should be the same; therefore, the Spearman-rho rank-order correlation between the TCI scores and insight should be significant and positive.

H3: CCT predicted that insight would be suppressed in more intuitive task conditions. This essentially should result in a truncation of range for insight scores as tasks become more intuitive. This in turn should result in an increasing discrepancy between insight and measures of performance; that is, less positive correlations between insight and measures of performance as tasks become more intuitive (i.e., have lower TCI scores). This can be expressed in the following manner: as tasks become more intuitive, the correlation between insight and performance indices becomes smaller and smaller. Simultaneously, the difference scores between each should become larger. The Spearman-rho rank-order correlations between TCI scores and the difference scores were predicted to be significant and negative.

H4: Because metacognition may be more properly conceived of as a *trait* as measured here (i.e., independently of task effects) and participants were randomly assigned to conditions, there was no predicted significant difference for metacognition scores by task condition. However, there was a predicted significant difference in the pattern of correlations between the metacognition scores and the performance indices.

Specifically, performance should be suppressed as tasks became more intuitive in orientation. However, as tasks became more analytical in nature, metacognition (and/or its proposed conceptual sibling insight) can be brought to bear and the correlation between metacognition and the performance indices should strengthen. This can be expressed in the following manner: as tasks become more intuitive, the correlation between metacognition and performance indices becomes smaller and smaller. Simultaneously, the difference scores between each should become larger. The Spearman-rho rank-order correlations between TCI scores and difference scores were predicted to be significant and negative.

H5: Finally, it was also predicted that the correlational pattern between insight and metacognition would follow the pattern of results predicted for the other independent variables. This reflected the nature of the “trait” vs “state” assignations that have been given to metacognition versus insight. Namely, metacognition was independent of the suppressing effects of intuitive tasks, at least as measured here. However, insight was not. As the task ecology allowed insight to exhibit a larger and larger presence (i.e., the tasks became more and more analytical), the similarity of the two sets of scores should have increased. Simultaneously, the difference scores between the two measures should decrease. The Spearman-rho rank-order correlation between TCI score and the difference scores were predicted to be significant and negative.

Methods

Design

The design was a between-subjects Task Type (analytical vs. intuitive) x Function Form (linear vs. quadratic) design. Three of the five dependent measures were the traditional Lens Model Equation performance indices of achievement, matching, and cognitive control. In addition, insight and metacognition were measured. The correlational patterns between the dependent variables were also analyzed.

Participants

Because the expected effect size was unknown, a power analysis based upon pilot data was conducted. A power level of .80 is usually considered reasonable (Chase & Tucker, 1976). Initial power analysis (based upon Cohen's *f* statistic) indicated that 20 individuals per cell were optimal. However, power analysis is usually used for designs in which individuals are observed once. Stability of measurement in the current design, which revolves around summary statistics calculated across a series of responses, achieves power with fewer participants.

Power analyses in such situations, therefore, are often treated as an upward boundary (Cooksey, 1996). Post-hoc power analyses indicated that n=15 in each cell were sufficient. Therefore, sixty college students from the University of Georgia Psychology research pool were recruited for this study. Students were given class credit for their participation.

Experimental Scenario

Participants first completed two measures of metacognition, the Metacognitive Awareness Inventory and a general knowledge test questionnaire consisting of multiple choice questions and confidence ratings (described above). The Metacognitive Awareness Inventory is a subjective self-questionnaire which measures engagement in behaviors associated with higher levels of metacognition.

All questions on the general knowledge test questionnaire asked for a comparison of population size between states in the U.S.A. All 75 of the questions were generated via a random pairing procedure, following the recommendations of Juslin and colleagues (Juslin, Winman, & Olsson, 2000). By calculating the absolute difference between average confidence and average performance (percent correct) across all 75 questions, an accuracy score for each participant was generated.

Participants were subsequently trained in one of four experimental scenarios (analytical-linear; analytical-quadratic; intuitive-linear; intuitive-quadratic). Successful learning of the task was indicated by a matching index that was positive and significantly different from zero. Participants were told that they were going to learn to assess the threat level of an incoming aircraft, based upon the cues shown to them onscreen. The correct answers to the first two blocks of trials would be supplied by a highly trained expert.

The cue-criterion function-form in the linear conditions was a positive function, and the cue-criterion function-form in the quadratic conditions was an inverted U-shape. Each participant was given 2 blocks of 60 training trials. Each trial was followed by immediate outcome feedback (the correct answer). The judgment analysis software used to present the information to the participants was POLICY PC.

At the end of each block of trials, each participant was also given cognitive feedback. Cognitive feedback consisted both of numerical information, such as the mean and standard deviation of their judgments as compared to those of the correct criterion values, as well as scatterplots showing both the overall function between cues and judgments they made, and the overall function between cues and criterion values. Finally, participants saw bar graphs showing both how much relative weight a cue was given in the ecology and in their judgments. Thus, participants saw if they were over or under weighting any given cue.

Following the training, participants completed a block of 80 more trials, absent any feedback. Once the block of trials was completed, participants filled out an insight measure requesting them to subjectively weight the cues as they had been used in the preceding block of trials. Participants also filled out a questionnaire assessing their levels of confidence in their performance (a.k.a. "answer confidence") versus confidence in their level of insight (a.k.a. "method confidence"). Finally, participants were asked to indicate how well they thought they had performed in the outcome-free block of trials.

Manipulation Check

Just as CCT provides a list of characteristics that allow for the a priori classification of tasks as more analytical or more intuitive (i.e., generation of TCI scores), so too does CCT provide some characteristics that should allow for the differentiation of performance arising from engagement in analytical or intuitive cognition.

One such characteristic is the difference between confidence in performance and confidence in insight. This questionnaire was used to generate difference scores. Greater confidence in performance than insight should be associated with more intuitive cognition (and hence poorer performance). Conversely, greater confidence in insight than performance should be correlated with more analytical cognition and hence higher levels of performance. That is, "(s)ince method confidence is expected to be high in analysis and answer confidence to be high in intuition, the greater the difference between these measures, the more analytic the subject's cognitive activity" (Hammond et al., 1987, p. 759).

Results

All results reported below are from the MANOVA. Descriptive statistics from this analysis are shown in Table 4.

Hypothesis 1

Contrary to predictions, achievement was higher in the linear than in the quadratic condition, $F(1, 56) = 121.43$, $p < .001$. Also contrary to prediction, achievement was higher in the intuitive than the analytical condition $F(1, 56) = 6.40$, $p < .01$. Also contrary to prediction, matching was higher in the linear than the quadratic conditions, $F(1, 56) = 121.38$, $p < .001$. There was no significant effect for task type. Also contrary to prediction, cognitive control was significantly higher in the linear than the quadratic conditions, $F(1, 56) = 632.95$, $p < .001$. Further, cognitive control was higher in the intuitive than the analytical conditions, $F(1, 56) = 10.63$, $p < .01$.

Hypothesis 2

Contrary to prediction, insight was higher in the linear than the quadratic conditions, $\underline{F}$ (1, 56) = 29.58, $\underline{p} < .001$. Also contrary to prediction, insight was higher in the intuitive than the analytical conditions, $\underline{F}$ (1, 56) = 12.65, $\underline{p} < .01$.

Correlational Hypotheses

To investigate the general correlational pattern between insight, metacognition, and the three remaining LME performance indices, the zero-order correlations across conditions are shown in Table 5. There were several things worthy of noting.

The first is that the proposed relationship between insight and the LME performance indices appeared to be supported. Insight was highly and positively correlated with matching, achievement, and cognitive control. Further, the objective metacognition measure (the accuracy scores) appeared to be positively and significantly related to both cognitive control and insight. Therefore, the thesis that insight and metacognition are similar constructs receives some support. However, this support had to be qualified because the subjective measure of the metacognition (the MAI scores) exhibited no significant correlations at all. Therefore, the difference scores for H3 through H5 were constructed utilizing only accuracy scores.

Hypothesis 3

Contrary to prediction, the Spearman correlation between the TCI scores and the insight-achievement difference scores was significant and positive, $\underline{r} = +.56$, $\underline{p} < .001$. The correlation between the TCI scores and the insight-matching difference scores was not significant. Finally, contrary to prediction, the correlation between the TCI scores and the insight-cognitive control difference scores was significant and positive, $\underline{r} = +.29$, $\underline{p} < .05$.

Hypothesis 4

The Spearman correlation between the TCI scores and the metacognition-achievement difference scores was not significant. Contrary to predictions, the correlation between the TCI scores and the metacognition-matching difference scores was significant and positive, $\underline{r} = +.31$, $\underline{p} < .05$. Finally, the correlation between the TCI scores and the metacognition-cognitive control difference scores was not significant.

Hypothesis 5

Finally, contrary to prediction, the Spearman correlation between TCI scores and metacognition-insight difference scores was significant and positive, $\underline{r} = +.25$, $\underline{p} < .01$.

Contributions of Insight and Metacognition to Performance

Although the correlations across all four conditions between the five dependent measures were informative, it was still considered useful to explore how such patterns varied from condition to condition. This was essentially the case because the correlational magnitudes were expected to vary from condition to condition. Because there were no significant interaction effects from the MANOVA, there was no rationale for examining the correlational patterns from cell to cell. Rather, the two levels of the function form IV were compared to each other in terms of correlational patterns, as were the two levels of the task type variable.

Correlational Patterns in Linear and Quadratic Conditions

The relevant correlational matrices for the linear and quadratic conditions are shown in Table 6. An interesting pattern emerged when examining the pattern of correlations for accuracy scores and insight (because MAI scores were not correlated in the omnibus correlation matrices, only the accuracy score measures were examined further). In the linear conditions, only insight predicted performance. However, in the quadratic conditions, both insight and metacognition appeared to predict performance. The accuracy scores were significantly correlated with two of the LME indices of performance, and very nearly so (p=.053) with the third. Likewise, insight was significantly correlated with two of the three LME indices. The extent to which insight and metacognition predicted unique variance in the quadratic conditions was assessed by computing two partial correlation matrices. Both are displayed in Table 7.

The first partial correlation matrix examined the correlations between accuracy scores and the three LME performance indices with insight partialled out. The second partial correlation matrix examined the correlations between insight and the three LME indices with the accuracy scores partialled out. It appeared that insight alone explained unique performance variance in the linear conditions, and that accuracy scores alone explained performance variance in the quadratic conditions.

However, the extent to which insight and metacognition jointly contributed to performance variance in the quadratic conditions was still to be assessed. As indicated in Table 7, both accuracy scores and insight were significantly correlated with achievement and cognitive control. Therefore, two simultaneous regression equations were computed with accuracy scores and insight as the predictors: one with achievement as the dependent variable, and one with cognitive control as the dependent variable. The results of these regressions are displayed in Table 8. In both instances, the accuracy scores were more predictive than insight.

Correlational Patterns in Intuitive and Analytical Tasks

Because task type yielded significant main effects in the MANOVA, the correlational patterns among the five dependent measures was also explored. The relevant correlational matrices are displayed in Table 9. Unlike the function form IV, there did not appear to be a different role played by metacognition and insight in the intuitive and analytical tasks. In both instances, insight alone appeared to predict performance.

Manipulation Check

An examination of the differential confidence scores revealed no significant difference between the intuitive and the analytical conditions, $\underline{F}$ (1, 58) = 3.28, $\underline{p}$ = .07. A closer examination of the difference scores revealed a (marginally non-significant) tendency for individuals in the intuitive conditions to express more confidence in insight than in performance. This pattern is, according to CCT, more characteristic of analytical cognition than intuitive.

However, an examination of the differential confidence scores did reveal a significant difference between the linear and quadratic conditions, $\underline{F}$ (1, 58) = 5.63, $\underline{p}$ < .05. The difference scores were significantly larger in the linear than the quadratic condition, indicating that individuals in the linear condition expressed more confidence in insight than in performance. In other words, contrary to CCT, the linear conditions triggered more analytical cognition than did the nonlinear conditions.

Experiment One Summary

In almost all cases, Cognitive Continuum Theory was contradicted. Performance indices tended to be higher in the linear than quadratic conditions, and the linear conditions triggered more analytical cognition than did the nonlinear conditions. Further, performance tended to be higher in the intuitive than the analytical conditions. There was also a nonsignificant tendency for analytical cognition to be engaged in slightly more often by persons in the intuitive rather than the analytical conditions.

The failure of task type (analytical versus intuitive) to generate effects consistent with CCT is puzzling, but there is one plausible hypothesis. CCT posits that predictability (discussed below) is just one of several equally important factors that differentiate intuitive from analytical tasks. Predictability, however, was not manipulated in the current study due to software limitations. (POLICY PC does not allow for the easy manipulation of predictability. When predictability was manipulated in the current experiment, POLICY PC generated inconsistent cognitive feedback regarding the relationship between cue and criterion values from trial block to trial block.)

The finding that performance was consistently higher in the linear rather than the quadratic conditions is also contrary to CCT. It also appeared that function form (quadratic vs. linear) had more of an impact than did task type. This can be shown in several different ways.

One is the observation that the function form manipulation resulted in a greater difference in modes of cognition than did task type, as indicated by the manipulation check results. Further, the average partial eta-squared value associated with the four function form effects was .49, compared to an average partial eta-squared value of .11 associated with the task type manipulation. That is, on average the function form manipulation explained approximately 5 times as much variance as did the task type manipulation. There was also some indication that metacognition and insight played different roles in the linear and quadratic tasks. It appeared that the only time that metacognition explained unique performance variance was in the quadratic conditions.

The finding that nonlinear cue-criterion relationships results in lower performance was more consistent with the findings of Brehmer and associates, who proposed a function form hierarchy hypothesis (see Appendix C and pp. 16 for further discussion). This hypothesis compares two types of linear functions to two types of nonlinear functions. Could the overall pattern of results from the first experiment be replicated, when the positive linear function was replaced with a negative one, and the inverted U-shaped quadratic function replaced with a U-shaped one? Finally, could the results from both experiments be mapped onto the function form hierarchy hypothesis of Brehmer et. al?

Table 3. TCI Scores by Experimental Condition.

		Function Form		
		Quadratic	Linear	Row Means
Task Type	Analytical	6.32	2.9	9.22
	Intuitive	5.06	1.4	6.46
	Column Means	11.38	2.15	

Table 4. Descriptive Statistics Experiment 1, H1-H2.

Descriptive Statistics

	Function Form	Task Type	Mean	Std. Deviation	N
achievement	linear	analytical	.8493	.29659	15
		intuitive	1.0747	.23712	15
		Total	.9620	.28765	30
	quadratic	analytical	.1819	.24343	16
		intuitive	.2900	.23716	14
		Total	.2323	.24264	30
	Total	analytical	.5048	.43087	31
		intuitive	.6959	.46202	29
		Total	.5972	.45273	60
matching	linear	analytical	1.6900	.52896	15
		intuitive	1.7013	.48879	15
		Total	1.6957	.50045	30
	quadratic	analytical	.3156	.43163	16
		intuitive	.4464	.38005	14
		Total	.3767	.40684	30
	Total	analytical	.9806	.84332	31
		intuitive	1.0955	.77058	29
		Total	1.0362	.80422	60
cognitive control	linear	analytical	.9913	.34424	15
		intuitive	1.3333	.38381	15
		Total	1.1623	.39821	30
	quadratic	analytical	.7450	.30509	16
		intuitive	.9200	.10806	14
		Total	.8267	.24752	30
	Total	analytical	.8642	.34278	31
		intuitive	1.1338	.35108	29
		Total	.9945	.36973	60
insight	linear	analytical	.8800	.31491	15
		intuitive	1.1260	.27856	15
		Total	1.0030	.31778	30
	quadratic	analytical	.5594	.19010	16
		intuitive	.7614	.15427	14
		Total	.6537	.19968	30
	Total	analytical	.7145	.30146	31
		intuitive	.9500	.29021	29
		Total	.8283	.31664	60

Table 5. Zero-order Correlations for Experiment 1.

Correlations

		achievement	matching	cognitive control	insight	MAI	Accuracy Scores
achievement	Pearson Correlation	1	.871**	.718**	.790**	.166	.220
	Sig. (2-tailed)	.	.000	.000	.000	.205	.091
	N	60	60	60	60	60	60
matching	Pearson Correlation	.871**	1	.416**	.559**	.114	.150
	Sig. (2-tailed)	.000	.	.001	.000	.385	.252
	N	60	60	60	60	60	60
cognitive control	Pearson Correlation	.718**	.416**	1	.791**	.133	.310*
	Sig. (2-tailed)	.000	.001	.	.000	.309	.016
	N	60	60	60	60	60	60
insight	Pearson Correlation	.790**	.559**	.791**	1	.126	.294*
	Sig. (2-tailed)	.000	.000	.000	.	.336	.023
	N	60	60	60	60	60	60
MAI	Pearson Correlation	.166	.114	.133	.126	1	.084
	Sig. (2-tailed)	.205	.385	.309	.336	.	.522
	N	60	60	60	60	60	60
Accuracy Scores	Pearson Correlation	.220	.150	.310*	.294*	.084	1
	Sig. (2-tailed)	.091	.252	.016	.023	.522	.
	N	60	60	60	60	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 6. Correlation Matrices for Linear and Quadratic Conditions, Experiment 1.

Correlations (Linear Conditions)

		achievement	matching	cognitive control	insight	accuracy scores
achievement	Pearson Correlation	1	.359	.944**	.885**	.282
	Sig. (2-tailed)	.	.051	.000	.000	.131
	N	30	30	30	30	30
matching	Pearson Correlation	.359	1	.083	.187	.128
	Sig. (2-tailed)	.051	.	.663	.322	.500
	N	30	30	30	30	30
cognitive control	Pearson Correlation	.944**	.083	1	.859**	.267
	Sig. (2-tailed)	.000	.663	.	.000	.154
	N	30	30	30	30	30
insight	Pearson Correlation	.885**	.187	.859**	1	.318
	Sig. (2-tailed)	.000	.322	.000	.	.087
	N	30	30	30	30	30
accuracy scores	Pearson Correlation	.282	.128	.267	.318	1
	Sig. (2-tailed)	.131	.500	.154	.087	.
	N	30	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations (Quadratic Conditions)

		cognitive control	matching	achievement	insight	Accuracy Scores
cognitive control	Pearson Correlation	1	.063	.163	.384*	.476**
	Sig. (2-tailed)	.	.740	.389	.036	.008
	N	30	30	30	30	30
matching	Pearson Correlation	.063	1	.971**	.266	.357
	Sig. (2-tailed)	.740	.	.000	.155	.053
	N	30	30	30	30	30
achievement	Pearson Correlation	.163	.971**	1	.363*	.398*
	Sig. (2-tailed)	.389	.000	.	.049	.029
	N	30	30	30	30	30
insight	Pearson Correlation	.384*	.266	.363*	1	.347
	Sig. (2-tailed)	.036	.155	.049	.	.060
	N	30	30	30	30	30
Accuracy Scores	Pearson Correlation	.476**	.357	.398*	.347	1
	Sig. (2-tailed)	.008	.053	.029	.060	.
	N	30	30	30	30	30

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Table 7. Partial Correlation Matrices for Quadratic Conditions, Experiment 1

Controlling for Insight

	CC	Match	Ach	AccScor
CC	1.0000	-.0439	.0276	.3960
	(0)	(27)	(27)	(27)
	P= .	P= .821	P= .887	P= .033
Match	-.0439	1.0000	.9737	.2925
	(27)	(0)	(27)	(27)
	P= .821	P= .	P= .000	P= .124
Ach	.0276	.9737	1.0000	.3117
	(27)	(27)	(0)	(27)
	P= .887	P= .000	P= .	P= .100
AccScor	.3960	.2925	.3117	1.0000
	(27)	(27)	(27)	(0)
	P= .033	P= .124	P= .100	P= .

Controlling for AccScor

	CC	Match	Ach	Insight
CC	1.0000	-.1299	-.0328	.2658
	(0)	(27)	(27)	(27)
	P= .	P= .502	P= .866	P= .163
Match	-.1299	1.0000	.9675	.1624
	(27)	(0)	(27)	(27)
	P= .502	P= .	P= .000	P= .400
Ach	-.0328	.9675	1.0000	.2613
	(27)	(27)	(0)	(27)
	P= .866	P= .000	P= .	P= .171
Insight	.2658	.1624	.2613	1.0000
	(27)	(27)	(27)	(0)
	P= .163	P= .400	P= .171	P= .

Table 8. Simultaneous Regression Equations in Quadratic Conditions, Experiment 1

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.369	2	.184	3.721	.037[a]
	Residual	1.338	27	.050		
	Total	1.707	29			

a. Predictors: (Constant), insight, Accuracy Scores

b. Dependent Variable: achievement

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.154	.187		.822	.419
	Accuracy Scores	1.275E-02	.007	.310	1.705	.100
	insight	.310	.221	.256	1.406	.171

a. Dependent Variable: achievement

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.500	2	.250	5.287	.012[a]
	Residual	1.277	27	.047		
	Total	1.777	29			

a. Predictors: (Constant), insight, Accuracy Scores

b. Dependent Variable: cognitive control

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.785	.183		4.287	.000
	Accuracy Scores	1.637E-02	.007	.390	2.241	.033
	insight	.309	.216	.249	1.433	.163

a. Dependent Variable: cognitive control

Table 9. Correlation Matrices for Analytical and Intuitive Conditions, Experiment 1.

Correlations (Intuitive)

		cognitive control	matching	achievement	insight	accuracy scores
cognitive control	Pearson Correlation	1	.466*	.800**	.813**	.221
	Sig. (2-tailed)	.	.011	.000	.000	.250
	N	29	29	29	29	29
matching	Pearson Correlation	.466*	1	.861**	.588**	.197
	Sig. (2-tailed)	.011	.	.000	.001	.306
	N	29	29	29	29	29
achievement	Pearson Correlation	.800**	.861**	1	.791**	.208
	Sig. (2-tailed)	.000	.000	.	.000	.278
	N	29	29	29	29	29
insight	Pearson Correlation	.813**	.588**	.791**	1	.318
	Sig. (2-tailed)	.000	.001	.000	.	.093
	N	29	29	29	29	29
Accuracy Scores	Pearson Correlation	.221	.197	.208	.318	1
	Sig. (2-tailed)	.250	.306	.278	.093	.
	N	29	29	29	29	29

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations (Analytical)

		cognitive control	matching	achievement	insight	Accuracy Scores
cognitive control	Pearson Correlation	1	.381*	.607**	.707**	.340
	Sig. (2-tailed)	.	.034	.000	.000	.062
	N	31	31	31	31	31
matching	Pearson Correlation	.381*	1	.899**	.566**	.104
	Sig. (2-tailed)	.034	.	.000	.001	.576
	N	31	31	31	31	31
achievement	Pearson Correlation	.607**	.899**	1	.779**	.195
	Sig. (2-tailed)	.000	.000	.	.000	.292
	N	31	31	31	31	31
insight	Pearson Correlation	.707**	.566**	.779**	1	.232
	Sig. (2-tailed)	.000	.001	.000	.	.210
	N	31	31	31	31	31
Accuracy Scores	Pearson Correlation	.340	.104	.195	.232	1
	Sig. (2-tailed)	.062	.576	.292	.210	.
	N	31	31	31	31	31

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

CHAPTER 4

EXPERIMENT TWO

Purposes and Rationale

The purposes of the second experiment were to 1) replicate and extend the findings of the first experiment and 2) examine the extent to which type of nonlinear and linear function forms affect performance.

Cue-Criterion Function Forms

The hierarchy of cue-criterion function forms discussed earlier consisted of two linear (positive, negative) and two nonlinear (U-shaped, inverted U-shaped) functions. The first experiment utilized a positive linear and an inverted U-shaped function form. The second experiment replaced the positive linear function with a negative linear function, and the inverted U-shaped function with a U-shaped function.

TCI Scores

One way of examining the extent to which type of linear and type of nonlinear function forms can impact decision making in MPCL environments, as well as to simultaneously highlight how CCT does not consider this in the construction of TCI scores, was to construct tasks with the same TCI score but different function form. Therefore, the linear tasks in Experiment 2 utilized the same cue sets as the linear tasks in Experiment 1. All other factors were held constant so that the TCI scores for the linear conditions were the same across both experiments. Similarly, the nonlinear tasks in

Experiment 2 utilized the same cue sets as the nonlinear tasks in Experiment 1, while holding all other factors constant so that the same TCI scores were assigned.

Hypotheses and Methods

The same 5 hypotheses assessed in the first experiment were examined in the second experiment. In addition, the correlational patterns between insight, metacognition, and the three Lens Model Equation performance indices were explored in the same fashion as before. In all respects except for function forms, the second experiment was identical to the first.

Results

All descriptive statistics from the MANOVA analysis are listed in Table 10.

Hypothesis 1

Contrary to predictions, achievement was significantly higher in the linear than the quadratic conditions, $\underline{F}$ (1, 56) = 63.01, $\underline{p}$ <.001. There was no significant effect for task type. Also contrary to prediction, matching was significantly higher in the linear than the quadratic conditions, $\underline{F}$ (1, 56) = 284.32, $\underline{p}$ < .001. There was no significant effect for task type. Finally, neither function form nor task type significantly impacted cognitive control.

Hypothesis 2

Contrary to prediction, insight was significantly higher in the linear than the quadratic conditions, $\underline{F}$ (1, 56) = 4.23, $\underline{p}$ < .05. There was no significant effect for task type.

Correlational Hypotheses

To investigate the general correlational pattern between insight, metacognition, and the three remaining LME performance indices, the zero-order correlations across conditions are shown in Table 11. There were several things worthy of noting.

The first was that the proposed relationship between insight and the LME performance indices appeared to be supported. Insight was significantly correlated with two of the three LME performance indices.

Further, the subjective metacognition measure (the MAI scores) appeared to be positively and significantly related to two of the three LME performance indices, as well as insight. Therefore, the thesis that insight and metacognition are similar constructs received some support. However, this support was qualified because the objective measure of metacognition (the accuracy scores) exhibited no significant correlations at all. Therefore, the difference scores for H3 through H5 were constructed utilizing only the MAI scores.

Hypothesis 3

The Spearman correlation between the insight-achievement difference scores and the TCI scores was not significant. Contrary to predictions, the Spearman correlation between the insight-matching difference scores and the TCI scores was significant and positive, $\underline{r} = +.52$, $\underline{p} < .001$. The Spearman correlation between the insight-cognitive control and the TCI scores was not significant.

Hypothesis 4

Contrary to prediction, the Spearman correlation between the metacognition-achievement difference scores and the TCI scores was significant and positive, $\underline{r} = +.36$, $\underline{p} < .01$. Also contrary to prediction, the Spearman correlation between the metacognition-matching difference scores and the TCI scores was significant and positive, $\underline{r} = +.53$, $\underline{p} <$

.001. Finally, the Spearman correlation between the metacognition-cognitive control and the TCI scores was not significant.

Hypothesis 5

The Spearman correlation between the metacognition-insight difference scores and the TCI scores was not significant.

Contributions of Insight and Metacognition to Performance

Because only the function form IV yielded significant differences with regards to the MANOVA results, only the linear and quadratic condition correlational patterns were compared to one another.

Correlational Patterns in Linear and Quadratic Conditions

The relevant correlational matrices for the linear and quadratic conditions are shown in Table 12. The same general pattern seen in Experiment 1 was replicated here. In the linear conditions, only insight predicted performance. However, in the quadratic conditions, both insight and metacognition appeared to predict performance. The extent to which insight and metacognition predicted unique variance in the quadratic conditions was assessed by computing two partial correlation matrices. Both are displayed in Table 13.

The first partial correlation matrix examined the correlations between the MAI scores and the three LME performance indices with insight partialled out. The second partial correlation matrix examined the correlations between insight and the three LME indices with the MAI scores partialled out. It appeared that insight alone explained unique performance variance in the linear conditions, but that both insight and MAI scores explained performance variance in the quadratic conditions.

However, the extent to which insight and metacognition jointly contribute to performance variance in the quadratic conditions had still to be assessed. As indicated in Table 11, both the MAI scores and insight were significantly correlated with achievement and matching. Therefore, two simultaneous regression equations were computed with the MAI scores and insight as the predictors: one with achievement as the dependent variable, and one with matching as the dependent variable. The results of these regressions are displayed in Table 14. In both instances, insight was more predictive than the MAI scores.

Manipulation Check

An examination of the differential confidence scores revealed no significant difference between the intuitive and the analytical conditions, F (1, 58) = 1.96, p = .10. Further, an examination of the differential confidence scores revealed no significant difference between the linear and quadratic conditions, F (1, 58) = 2.05, p > .05.

Experiment Two Summary

In general, the second experiment was a successful replication of the first. The CCT predictions were overwhelmingly rejected. The linear conditions displayed significantly higher values for three of the four traditional LME performance indices. Further, it appeared that, once again, metacognition explained unique variance only in the quadratic conditions. Conversely, in the linear conditions insight alone explained performance variance. Once again, the function form manipulation appeared to be driving the differences observed. In all four cases, the function form manipulation effects were highly significant, compared with no significant effects for the task type manipulation. In addition, the average partial eta-squared value for the function form manipulation was approximately ten times as large as the average partial eta-squared value for task type (.31 and .03, respectively).

The findings were thus consistent those of the first experiment. However, the extent to which the function forms effects seen in both experiments mapped onto the function form hierarchy hypothesis remained to be explored.

Table 10. Descriptive Statistics for Experiment Two, H1-H2.

Descriptive Statistics

	Function Form	Task Type	Mean	Std. Deviation	N
Achievement	linear	analytical	.8333	.22906	15
		intuitive	.9847	.37395	15
		Total	.9090	.31426	30
	quadratic	analytical	.3713	.29181	15
		intuitive	.2493	.25260	15
		Total	.3103	.27525	30
	Total	analytical	.6023	.34877	30
		intuitive	.6170	.48801	30
		Total	.6097	.42060	60
matching	linear	analytical	1.7467	.66629	15
		intuitive	1.6800	.46098	15
		Total	1.7133	.56396	30
	quadratic	analytical	.6400	.48206	15
		intuitive	.4047	.40736	15
		Total	.5223	.45455	30
	Total	analytical	1.1933	.80202	30
		intuitive	1.0423	.77675	30
		Total	1.1178	.78646	60
cognitive control	linear	analytical	.9907	.28599	15
		intuitive	1.1387	.37707	15
		Total	1.0647	.33733	30
	quadratic	analytical	.8093	.57517	15
		intuitive	.9673	.24218	15
		Total	.8883	.44099	30
	Total	analytical	.9000	.45574	30
		intuitive	1.0530	.32334	30
		Total	.9765	.39928	60
insight	linear	analytical	1.6253	.89521	15
		intuitive	1.7220	.58536	15
		Total	1.6737	.74479	30
	quadratic	analytical	1.6707	.58031	15
		intuitive	.9907	.43559	15
		Total	1.3307	.61136	30
	Total	analytical	1.6480	.74161	30
		intuitive	1.3563	.62876	30
		Total	1.5022	.69734	60

Table 11. Zero-order Correlations for Experiment 2.

Correlations

		cognitive control	matching	achievement	MAI	insight	accuracy scores
cognitive control	Pearson Correlation	1	.310*	.595**	.063	.095	.026
	Sig. (2-tailed)	.	.016	.000	.630	.469	.842
	N	60	60	60	60	60	60
matching	Pearson Correlation	.310*	1	.892**	.303*	.576**	-.018
	Sig. (2-tailed)	.016	.	.000	.018	.000	.892
	N	60	60	60	60	60	60
achievement	Pearson Correlation	.595**	.892**	1	.295*	.485**	-.100
	Sig. (2-tailed)	.000	.000	.	.022	.000	.445
	N	60	60	60	60	60	60
MAI	Pearson Correlation	.063	.303*	.295*	1	.381**	.168
	Sig. (2-tailed)	.630	.018	.022	.	.003	.199
	N	60	60	60	60	60	60
insight	Pearson Correlation	.095	.576**	.485**	.381**	1	-.103
	Sig. (2-tailed)	.469	.000	.000	.003	.	.434
	N	60	60	60	60	60	60
accuracy scores	Pearson Correlation	.026	-.018	-.100	.168	-.103	1
	Sig. (2-tailed)	.842	.892	.445	.199	.434	.
	N	60	60	60	60	60	60

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Table 12. Correlation Matrices for Linear and Quadratic Conditions, Experiment 2.

Correlations (Linear Conditions)

		cognitive control	matching	achievement	MAI	insight
cognitive control	Pearson Correlation	1	.220	.808**	.074	-.155
	Sig. (2-tailed)	.	.243	.000	.699	.415
	N	30	30	30	30	30
matching	Pearson Correlation	.220	1	.660**	.302	.622**
	Sig. (2-tailed)	.243	.	.000	.104	.000
	N	30	30	30	30	30
achievement	Pearson Correlation	.808**	.660**	1	.331	.341
	Sig. (2-tailed)	.000	.000	.	.074	.065
	N	30	30	30	30	30
MAI	Pearson Correlation	.074	.302	.331	1	.480**
	Sig. (2-tailed)	.699	.104	.074	.	.007
	N	30	30	30	30	30
insight	Pearson Correlation	-.155	.622**	.341	.480**	1
	Sig. (2-tailed)	.415	.000	.065	.007	.
	N	30	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations (Quadratic Conditions)

		cognitive control	matching	achievement	MAI	insight
cognitive control	Pearson Correlation	1	.237	.518**	.049	.228
	Sig. (2-tailed)	.	.208	.003	.798	.225
	N	30	30	30	30	30
matching	Pearson Correlation	.237	1	.914**	.564**	.614**
	Sig. (2-tailed)	.208	.	.000	.001	.000
	N	30	30	30	30	30
achievement	Pearson Correlation	.518**	.914**	1	.453*	.615**
	Sig. (2-tailed)	.003	.000	.	.012	.000
	N	30	30	30	30	30
MAI	Pearson Correlation	.049	.564**	.453*	1	.356
	Sig. (2-tailed)	.798	.001	.012	.	.054
	N	30	30	30	30	30
insight	Pearson Correlation	.228	.614**	.615**	.356	1
	Sig. (2-tailed)	.225	.000	.000	.054	.
	N	30	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 13. Partial Correlation Matrices for Quadratic Conditions, Exp 2.

Controlling for Insight

	CC	Match	Ach	MAI
CC	1.0000	.1256	.4925	-.0356
	(0)	(27)	(27)	(27)
	P= .	P= .516	P= .007	P= .855
Match	.1256	1.0000	.8619	.4687
	(27)	(0)	(27)	(27)
	P= .516	P= .	P= .000	P= .010
Ach	.4925	.8619	1.0000	.3176
	(27)	(27)	(0)	(27)
	P= .007	P= .000	P= .	P= .093
MAI	-.0356	.4687	.3176	1.0000
	(27)	(27)	(27)	(0)
	P= .855	P= .010	P= .093	P= .

Controlling for MAI

	CC	Match	Ach	Insight
CC	1.0000	.2538	.5573	.2261
	(0)	(27)	(27)	(27)
	P= .	P= .184	P= .002	P= .238
Match	.2538	1.0000	.8947	.5362
	(27)	(0)	(27)	(27)
	P= .184	P= .	P= .000	P= .003
Ach	.5573	.8947	1.0000	.5453
	(27)	(27)	(0)	(27)
	P= .002	P= .000	P= .	P= .002
Insight	.2261	.5362	.5453	1.0000
	(27)	(27)	(27)	(0)
	P= .238	P= .003	P= .002	P= .

Table 14. Simultaneous Regression Equations in Quadratic Conditions, Experiment 2.

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3.082	2	1.541	14.295	.000[a]
	Residual	2.910	27	.108		
	Total	5.992	29			

a. Predictors: (Constant), Insight, MAI

b. Dependent Variable: matching

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.826	.305		-2.704	.012
	MAI	.182	.066	.396	2.757	.010
	Insight	.352	.107	.474	3.301	.003

a. Dependent Variable: matching

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.970	2	.485	10.670	.000[a]
	Residual	1.227	27	.045		
	Total	2.197	29			

a. Predictors: (Constant), Insight, MAI

b. Dependent Variable: achievement

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.362	.198		-1.824	.079
	MAI	7.477E-02	.043	.268	1.740	.093
	Insight	.234	.069	.520	3.380	.002

a. Dependent Variable: achievement

CHAPTER 5

GENERAL DISCUSSION

Function Forms

Because there were no significant interactions between the task type and the function form independent variables in either experiment, the data from both experiments were collapsed across task type so that just the effects of function form were examined. This was done to allow for the examination of the extent to which the results could be explained via the function form hierarchy hypothesis. The results from this MANOVA are displayed in Table 15. All significant differences were explored via Bonferroni t-tests, the results of which are displayed in Table 16.

Lens Model Equation Performance Indices

It appeared that the level of insight was significantly different between the kinds of nonlinear tasks. Similarly, the level of insight was significantly different in the types of linear tasks. Further, there was at least one instance in which the type of linear and nonlinear functions chosen when examining cognitive control could affect the outcome. For instance, the difference between the levels of cognitive control in a positive linear task and a U-shaped function was statistically significant. However, this was not the case when comparing cognitive control levels between a U-shaped function and a negative linear task.

This pattern of results would be expected if the function form hierarchy hypothesis was true, because the positive linear and U-shaped function forms are the first and third functions in the hierarchy, and are thus farther apart than the negative linear and U-shaped function forms, which are the second and third, respectively. This rationale is also consistent with the observation that the level of achievement in an inverted U-shape quadratic task was not significantly different from the U-shape, but became increasingly significant as one approached the positive linear function. This pattern indicates a shortcoming of CCT, which considers only linearity and nonlinearity of function form.

A clearer indication of how the findings map onto the function form hierarchy hypothesis would be indicated by a trend analysis. Unfortunately, an F-test of linear trend was inappropriate because the assumptions of the test (quantitative variable, evenly spaced intervals) were not met. However, a rough indication of the extent to which the results for all four LME performance indices corresponded with placement in the function form hierarchy was given by assigning values of 1 to 4 to the inverted U-shape, U-shape, negative linear, and positive linear function forms, respectively.

The resulting correlations are shown in Table 17, and bar graphs of the relationships underlying these correlations are shown in Figure 2. In all four cases, the Spearman correlations between hierarchy position and the four LME performance indices (accuracy, matching, cognitive control, insight) were positive and highly significant. These results are consistent with the hierarchy hypothesis.

Manipulation Check Difference Scores

CCT predicted that higher levels of performance should be associated with greater confidence in insight than performance. Although CCT incorrectly predicted which conditions would result in greater difference scores, it was the case that higher performance covaried positively with the magnitude of the difference scores.

Interestingly, however, the pattern of the confidence scores was also consistent with the hierarchy hypothesis.

The hierarchy hypothesis would predict that the differences observed in the first experiment, which contrasted the first and fourth position members, should result in overall greater differences than the second experiment, which contrasted the second and the third position members. This pattern did emerge in that the confidence scores in the first experiment were significantly different between the linear and quadratic conditions, but not so (although in the correct direction, i.e., linear higher than quadratic) in the second experiment.

Effect Size Estimates

The hierarchy hypothesis would also predict that the effect sizes for function form in the first experiment would be larger than those for function form in the second experiment. This rests upon the fact that the first experiment contrasted the first and fourth members of the hierarchy, whereas the second contrasted the second and third. Results were consistent with this prediction, with an average partial eta-squared value for function form in the first experiment of .49, as contrasted with an average partial eta-squared value of .31 for function form in the second experiment.

Metacognition, Insight, and Performance

The manner in which metacognition and insight contributed to performance in the linear and quadratic conditions could, it is argued, also have been seen as moving along a continuum paralleling that of the function form hierarchy. In the positive linear condition, insight alone explained unique performance variance. In the negative linear condition, once again only insight explained performance variance. In the U-shaped function form, both insight and metacognition explained unique performance variance. However, in the inverted U-shaped function form, metacognition alone explained unique variance. It appeared that as tasks became more difficult, metacognition became more salient, and insight less so.

The position that metacognition and insight are similar constructs was somewhat supported. As noted above, there appeared to be a “dissociation” between metacognition and insight. As tasks became more difficult (e.g., were placed higher in the the function form hierarchy), metacognition became more predictive of performance, and insight became less so. Further, as tasks became more difficult, the correlation between metacognition and insight decreased.

Task Type

The task variable exhibited three significant effects in the first experiment, but none in the second. Thus, there appeared to be mixed evidence concerning the extent to which task type drives performance, at least in the absence of lowered predictability. The present evidence appeared to suggest that when task type significantly impacted performance, intuitive tasks led to higher performance than analytical tasks. This is contrary to CCT, which predicts higher performance in analytical tasks. One possible explanation for this finding revolves around the concept of predictability.

As mentioned above, a commonly manipulated factor in CCT research is lower predictability for intuitive environments. However, in the assignation of Task Continuum Index (TCI) scores, CCT places no more emphasis upon predictability than it does the other factors listed in Table 1. Therefore, according to CCT it is possible to make up for similarity of predictability by increasing the discrepancy between the other task factors.

However, MCPL research indicates that predictability is strongly and positively related to performance (Dudycha & Naylor, 1966; Brehmer, 1974b).

The fact that the pattern of results normally seen in CCT is reversed when predictability is not manipulated suggests that the driving force behind the differing performance profiles of intuitive and analytical tasks may be largely due to the lower predictability of intuitive tasks. This hypothesis is indirectly supported by the observation that the pilot study, which *did* manipulate predictability, yielded results consistent with CCT (i.e., positive correlations between performance indices and the TCI scores; significantly greater performance in the nonlinear than linear conditions; significantly greater performance in the analytical than the intuitive tasks).

Limitations and Future Experiments

One important limitation of the current experiment concerns predictability. Due to the nature of the software, predictability was not manipulated. As indicated, however, predictability may in fact be more important a factor than acknowledged by CCT. One manner of examining whether or not predictability is the main factor driving the lower performance usually seen in intuitive conditions would be to construct analytical and intuitive tasks which varied in a similar fashion along a predictability continuum (e.g., both types of task at .20, .40, .60, and .80 levels of predictability). If predictability is indeed the major factor, it should be possible to show that more predictable analytical tasks result in higher performance levels than do less predictable intuitive tasks. More convincingly, it could also be shown that more predictable intuitive tasks exhibit higher performance indices than do less predictable analytical tasks.

The finding that metacognition was more predictive of performance in tasks of greater difficulty is intriguing. There are two possible explanations for this finding. The first is simply that when tasks are easier, it is not necessary to possess high levels of metacognition to do well. The more difficult the task becomes, the more dependent performance is upon metacognition. If the relationship between metacognition and performance is indeed driven by task difficulty, there are many avenues of exploration suggested.[2] For example, lowerering predictability (i.e., greater difficulty, a. k. a. the "hard" effect) might also be expected to increase the correlation between metacognition and performance. Imposing increasingly stringent time constraints upon the task might serve the same purpose.

The second possible explanation involves a shortcoming of POLICY PC. POLICY PC is a software package that is self-paced in nature, and does not allow for the measurement of response rate. Therefore, it is possible that more metacognitive individuals would, when completing a more difficult task, engage in greater amounts of monitoring behavior (pondering an answer, spending more time assessing strategies, etc.).

If this were the case, then two things should be observed. Firstly, a positive correlation should be seen between the amount of time spent on a block of trials and measures of metacognition. Secondly, imposing increasingly stringent time constraints would, presumably, attenuate this correlation if the relationship is in fact due to the application of time and resource demanding strategies. This is rather elegant, in that the

[2] This is not the same argument underlying the hard-easy effect. In general, the hard-easy effect measures discrepancies between performance and confidence ratings elicited *within* a task. The current argument concerns itself with the relationship between metacognition measured independent of a task. Further, the application of metacognitive measures to MCPL tasks is fairly new.

same manipulation (increasing time constraints) should lead to one outcome or the other depending upon which explanation is correct. Further, it would be of interest to see if the correlation between metacognition and performance holds true when there is little time for application of resource-demanding strategies.

Finally, the observation that there was an inconsistent relationship between the two measures of metacognition and insight is problematic. In the first experiment, only the accuracy scores were correlated with performance and insight. In the second experiment, only the MAI scores were correlated with performance and insight. When the zero-order correlations between performance and the two measures of metacognition are examined across the pooled data from both experiments[3] (see Table 21), the MAI scores alone predicted performance and insight (although the correlation between achievement and accuracy scores approaches significance, $\underline{p} = .07$). Further, the metacognitive scores were themselves not correlated with each other.

Taken together, these two findings suggest that the inconsistent pattern may be due to the fact that insight is more correlated with some aspects of metacognition than others. That is, the metacognitive measures may not have been correlated with each other because they were measuring different aspects of the same construct. Further, it is plausible that insight correlates more highly with some aspects of metacognition than others. This is sensible because the foundation underlying the argument for conceptual similarity between insight and metacognition rested upon the conceptualization of insight as a measure of declarative knowledge. Declarative knowledge, however, is only one of

[3] Although there is little doubt that the four function forms utilized in the current project do not exhaust the universe of possibilities, there is ample evidence that the four function forms do approximate a meaningful continuum in terms of what humans can learn. It has been found, in other words, that utilizing more complex predictor terms than simple quadratic ones in computing policies adds little in terms of explanatory power (Cooksey, 1996).

several important aspects of metacognition. In other words, it may be that measures of declarative knowledge would correlate more highly with insight than other measures of metacognition.

There is, however, another possible explanation for the correlations seen between insight and metacognition. This is known as methods variance. Essentially, methods variance is an acknowledgment that the manner in which constructs are measured can sometimes inflate or attenuate a correlation between the constructs. For example, the correlations between similar constructs are sometimes greater when both are measured in a similar fashion (e.g., paper and pencil) than when measured in different fashions (e.g., paper and pencil versus a computer administered questionnaire). In more extreme cases, it is possible that similar methods of measurement can result in significant correlations between different constructs.

This speculation is germane to the current paper because insight and metacognition were both measured via paper and pencil questionnaires. [4] Therefore, a fruitful avenue of exploration would seek to answer both questions: to what extent does insight correlate with various aspects of metacognition, and how are those correlations affected by method of measurement?

Both questions indicate the potential usefulness of an approach known as multitrait-multimethod (MTMM; Campbell & Fiske, 1959). MTMM allows for the examination of both traits (e.g., insight and metacognition in its various aspects) as well as methods of measurement (paper and pencil, verbalizations, computer based questionnaires, etc.) as sources of variance. Measuring different aspects of metacognition

---

[4] However, the judgment task was administered via computer. Therefore, it could also be argued that the correlations seen between insight and performance (as well as those between metacognition and performance) might have been *lessened* due to methods variance.

via several methods, as well as insight via the same methods, would help illuminate the possibilities discussed above.

Table 15. MANOVA Results for Hierarchy Position Hypothesis

Univariate Tests

Dependent Variable		Sum of Squares	df	Mean Square	F	Sig.
cognitive control	Contrast	2.166	3	.722	5.469	.001
	Error	15.315	116	.132		
matching	Contrast	47.574	3	15.858	67.435	.000
	Error	27.278	116	.235		
achievement	Contrast	13.367	3	4.456	56.376	.000
	Error	9.168	116	7.903E-02		
insight	Contrast	17.217	3	5.739	21.467	.000
	Error	31.011	116	.267		

The F tests the effect of hierarchy position. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

Table 16. Post-hoc Bonferroni Tests for Hierarchy Position Hypothesis.

Multiple Comparisons

Bonferroni

Dependent Variable	(I) combo	(J) combo	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
cognitive control Fisher's r to z	pos lin	neg lin	.0977	.09382	1.000	-.1542	.3495
		U-shape	.2740*	.09382	.025	.0222	.5258
		Inv U shape	.3357*	.09382	.003	.0838	.5875
	neg lin	pos lin	-.0977	.09382	1.000	-.3495	.1542
		U-shape	.1763	.09382	.376	-.0755	.4282
		Inv U shape	.2380	.09382	.075	-.0138	.4898
	U-shape	pos lin	-.2740*	.09382	.025	-.5258	-.0222
		neg lin	-.1763	.09382	.376	-.4282	.0755
		Inv U shape	.0617	.09382	1.000	-.1902	.3135
	Inv U shape	pos lin	-.3357*	.09382	.003	-.5875	-.0838
		neg lin	-.2380	.09382	.075	-.4898	.0138
		U-shape	-.0617	.09382	1.000	-.3135	.1902
matching Fisher's r to z	pos lin	neg lin	-.0177	.12521	1.000	-.3538	.3184
		U-shape	1.1733*	.12521	.000	.8372	1.5094
		Inv U shape	1.3190*	.12521	.000	.9829	1.6551
	neg lin	pos lin	.0177	.12521	1.000	-.3184	.3538
		U-shape	1.1910*	.12521	.000	.8549	1.5271
		Inv U shape	1.3367*	.12521	.000	1.0006	1.6728
	U-shape	pos lin	-1.1733*	.12521	.000	-1.5094	-.8372
		neg lin	-1.1910*	.12521	.000	-1.5271	-.8549
		Inv U shape	.1457	.12521	1.000	-.1904	.4818
	Inv U shape	pos lin	-1.3190*	.12521	.000	-1.6551	-.9829
		neg lin	-1.3367*	.12521	.000	-1.6728	-1.0006
		U-shape	-.1457	.12521	1.000	-.4818	.1904
achievement Fisher's r to z	pos lin	neg lin	.0530	.07259	1.000	-.1418	.2478
		U-shape	.6517*	.07259	.000	.4568	.8465
		Inv U shape	.7297*	.07259	.000	.5348	.9245
	neg lin	pos lin	-.0530	.07259	1.000	-.2478	.1418
		U-shape	.5987*	.07259	.000	.4038	.7935
		Inv U shape	.6767*	.07259	.000	.4818	.8715
	U-shape	pos lin	-.6517*	.07259	.000	-.8465	-.4568
		neg lin	-.5987*	.07259	.000	-.7935	-.4038
		Inv U shape	.0780	.07259	1.000	-.1168	.2728
	Inv U shape	pos lin	-.7297*	.07259	.000	-.9245	-.5348
		neg lin	-.6767*	.07259	.000	-.8715	-.4818
		U-shape	-.0780	.07259	1.000	-.2728	.1168
insight_match_r	pos lin	neg lin	-.6707*	.13350	.000	-1.0290	-.3123
		U-shape	-.3277	.13350	.094	-.6860	.0307
		Inv U shape	.3493	.13350	.060	-.0090	.7077
	neg lin	pos lin	.6707*	.13350	.000	.3123	1.0290
		U-shape	.3430	.13350	.069	-.0153	.7013
		Inv U shape	1.0200*	.13350	.000	.6617	1.3783
	U-shape	pos lin	.3277	.13350	.094	-.0307	.6860
		neg lin	-.3430	.13350	.069	-.7013	.0153
		Inv U shape	.6770*	.13350	.000	.3187	1.0353
	Inv U shape	pos lin	-.3493	.13350	.060	-.7077	.0090
		neg lin	-1.0200*	.13350	.000	-1.3783	-.6617
		U-shape	-.6770*	.13350	.000	-1.0353	-.3187

Based on observed means.

*. The mean difference is significant at the .05 level.

Table 17. Spearman Correlations Between Hierarchy Position and Performance.

	Cognitive Control	Matching	Achievement	Insight
Hierarchy Position	.36***	.75***	.74***	.33***

***Correlation is significant at the .001 level.

Table 18. Zero-order Correlations Across Experiments 1 and 2.

Correlations

		cognitive control	matching	achievement	MAI	insight	accuracy scores
cognitive control	Pearson Correlation	1	.359**	.654**	.096	.230*	.165
	Sig. (2-tailed)	.	.000	.000	.295	.012	.071
	N	120	120	120	120	120	120
matching	Pearson Correlation	.359**	1	.880**	.209*	.478**	.070
	Sig. (2-tailed)	.000	.	.000	.022	.000	.448
	N	120	120	120	120	120	120
achievement	Pearson Correlation	.654**	.880**	1	.229*	.465**	.070
	Sig. (2-tailed)	.000	.000	.	.012	.000	.448
	N	120	120	120	120	120	120
MAI	Pearson Correlation	.096	.209*	.229*	1	.244**	.041
	Sig. (2-tailed)	.295	.022	.012	.	.007	.659
	N	120	120	120	120	120	120
insight	Pearson Correlation	.230*	.478**	.465**	.244**	1	.029
	Sig. (2-tailed)	.012	.000	.000	.007	.	.754
	N	120	120	120	120	120	120
accuracy scores	Pearson Correlation	.165	.070	.070	.041	.029	1
	Sig. (2-tailed)	.071	.448	.448	.659	.754	.
	N	120	120	120	120	120	120

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

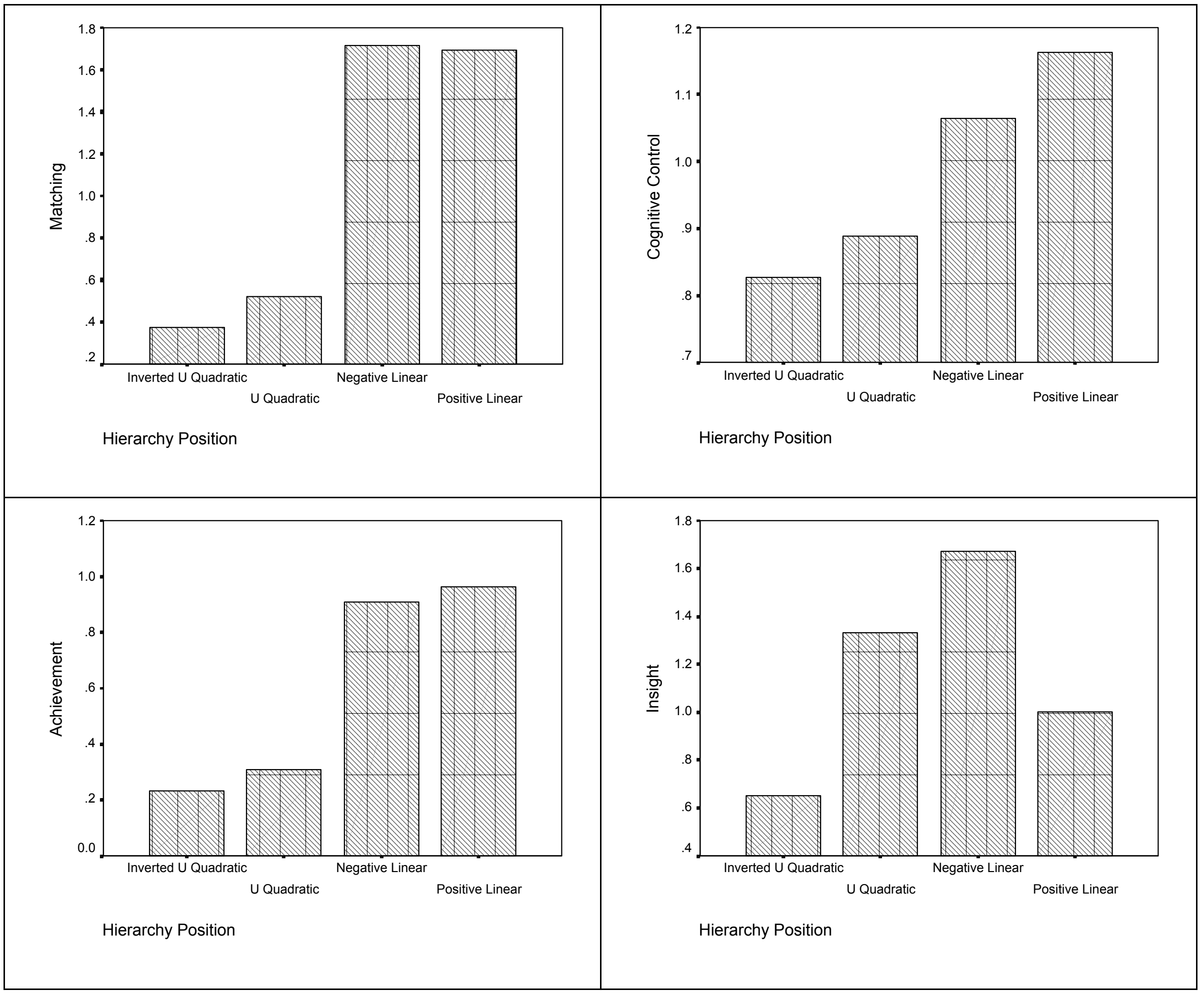

Figure 2. Function Form Hierarchy Relationships

REFERENCES

Ahl, V. A., Moore, C. F., & Dixon, J. A. (1992). Development of intuitive and numerical proportional reasoning. *Cognitive Development, 7,* 81-108.

Armelius, B. A., & Armelius, K. (1976). Confidence and performance in probabilistic inference tasks with interrelated cues. *Umea Psychological Reports* (96).

Armelius, K., & Armelius, B. A. (1975). Confidence in multiple-cue judgments as a function of cue intercorrelation and task predictability. *Umea Psychological Reports* (82).

Armelius, B. A., & Armelius, K. (1975). Integration rules in a multiple-cue probability learning task with intercorrelated cues. *Umea Psychological Reports* (80).

Artzt, A. F., & Armour-Thomas, E. (1996). Mathematics teaching as problem solving: A framework for studying teacher metacognition underlying instructional practice in mathematics. *Instructional Science, 26* (1-2), 5-25.

Balzer, W. K., Doherty, M. E., & O'Connor, R. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin, 106* (3), 410-433.

Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *Academy Management Review, 3*, 439-449.

Bisanz, G. L., Vesonder, G. T., & Voss, J. F. (1978). Knowledge of one's own responding and the relation of such knowledge to learning: A developmental study. *Journal of Experimental Child Psychology, 25* (1), 116-128.

Björkman, M. (1965). Learning of linear functions: Comparison between a positive and a negative slope. *University of Stockholm, Psychological Laboratories* (Report 183).

Bowman, E. H. (1963). Consistency and optimality in management decision making. *Management Science, 9*, 310-321.

Brehmer, B. (1988). The development of social judgment theory. In B. Brehmer & C. R. B. Joyce (Eds.), *Human judgment: The social judgment theory view* (pp. 13-40). Amsterdam: North-Holland.

Brehmer, B. (1974a). Hypotheses about relations between scaled variables in the learning of probabilistic inference tasks. *Organizational Behavior and Human Performance, 11*, 1-17.

Brehmer, B. (1974b). The effect of cue intercorrelation on interpersonal learning of probabilistic inference tasks. Organizational Behavior and Human Decision Processes, 12(3), 397-412.

Brehmer, B. (1977). Cognitive skills and optimal performance in judgment tasks. *Umea Psychological Reports* (107).

Brehmer, B. (1976). Testing hypotheses about functional relations in probabilistic inference tasks. *Umea Psychological Reports* (89).

Brehmer, B., & Svensson, C. (1976). Learning to use functional rules in inference tasks. Scandinavian Journal of Psychology, 17(4), 313-319.

Brehmer, B., & Qvarnstrom, G. (1973). Feedforward, feedback, and subjective weights in multiple-cue judgments. *Umea Psychological Reports* (76).

Brehmer, B. (1969). Cognitive dependence on additive and configural cue-criterion relations. *American Journal of Psychology, 82* (4), 490-503.

Brehmer, B. (1971). Subjects' ability to use functional rules. *Psychonomic Science, 24* (6), 259-260.

Brooks, L. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. Lloyd, *Cognition and Categorization* (pp. 170-211).

Brunswik, E. (1956a). Historical and thematic relations of psychology to other sciences. *Scientific Monthly, 83,* 151-161.

Brunswik, E. (1956b). Perception and thinking. In K. R. Hammond & T. R. Stewart (Eds.), *The Essential Brunswik* (pp. 260-264).

Brunswik, E. (1939). Probability as a determiner of rat behavior. In K. R. Hammond & T. R. Stewart (Eds.), *The Essential Brunswik* (pp. 195-210).

Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological Review, 50*, 255-272.

Brunswik, E. (1952). *The conceptual framework of psychology*. Chicago: University of Chicago Press.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*, 245-282.

Byrne, R. (1995). *The thinking ape: Evolutionary origins of intelligence*. New York: Oxford University Press.

Camerer, C. F. (1981a). <u>The validity and utility of expert judgment.</u> Unpublished doctoral dissertation, Center for Decision Research, University of Chicago. Graduate School of Business, Chicago.

Camerer, C. F. (1981b). General conditions for the success of bootstrapped models. *Organizational Behavior and Human Performance, 27*, 411-422.

Camerer, C. F., & Johnson, E. J. (1997). The process-performance paradox: How can experts know so much and predict so badly? In W. M. Goldstein & R. M. Hogarth (Eds.), *Research on Judgment and Decision Making* (pp. 342-364). Cambridge, UK: Cambridge University Press.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Chase, L. J., & Tucker, R. K. (1976). Statistical power: Derivation, development, and data-analytic implications. *The Psychological Record, 26*, 473-486.

Coleman, E. B., & Shore, B. (1991). Problem-solving processes of high and average performers in physics. *Journal for the Education of the Gifted, 14* (4), 366-379.

Cook, J. K. (1967). MMPI in actuarial diagnoses of psychological disturbance among college males. *Journal of Counseling Psychology, 14* (5), 474-477.

Cooksey, R. W. (1996). *Judgment analysis: Theory, methods, and applications.* San Diego: Academic Press.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81*, 95-106.

de Klerk, L. F., & Vroon, A. G. (1974). The effect of cue-criterion function form on single-cue learning. *Journal of Experimental Psychology, 102* (5), 890-892.

de Klerk, L. F., & Oppe, S. (1972). Functional cue learning under different probabilistic feedback conditions. *Psychonomic Science, 29* (6-B), 377-380.

Doherty, M. E. (2001). While God may not gamble, animals and humans do.

In K. R. Hammond & T. R. Stewart (Eds.), *The Essential Brunswik* (pp. 192-195).

Doherty, M. E., Brake, G. L., & Kleiter, G. D. (2001). The contribution of representative design to calibration research. In K. R. Hammond & T. R. Stewart (Eds.), *The Essential Brunswik* (pp. 317-320).

Doherty, M. E., & Balzer, W. K. (1988). Cognitive feedback. In B. Brehmer & C. R. B. Joyce (Eds.), *Human judgment: The SJT view. Advances in psychology* (pp. 163-197): Amsterdam, Netherlands: North Holland.

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

Dunwoody, P. T. (1998). Cognitive adaptation and its consequences: A test of Cognitive Continuum Theory. Unpublished master's thesis, University of Georgia, Athens, GA.

Earle, T. C. (1973). Intuitive and analytical thinking in consistent and inconsistent multiple-cue learning tasks. Unpublished doctoral dissertation, University of Colorado, Department of Psychology.

Einhorn, H. J., Kleinmuntz, D. N., & Kleinmuntz, B. (1979). Linear regression and process-tracing models of judgment. *Psychological Review, 86*, 465-485.

Eisler, H., & Spolander, K. (1970). On the sign of slope in the learning of linear functions. *Scandinavian Journal of Psychology, 11* (3), 176-184.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: The MIT Press.

Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cognitive Psychology, 1* (4), 324-340.

Flavell, J. H. (1979). Metacognition and cognitive monitoring. *American Psychologist, 34* (10), 906-911.

Gourgey, A. F. (1998). Metacognition in basic skills instruction. *Instructional Science, 26*, 81-96.

Gray, C. W. (1979). Ingredients of intuitive regression. *Organizational Behavior and Human Decision Processes, 23*(1), 30-48.

Haarbauer. E. S. (1996). <u>The Insight Continuum Index: Self-insight as a Measure of Task Properties in Cognitive Continuum Theory</u>. Unpublished master's thesis, University of Georgia, Athens, GA.

Hammond, K. R. (1955). Probabilistic functionalism and the clinical method. *Psychological Review, 71*, 438-456.

Hammond, K. R. (1966). *The Psychology of Egon Brunswik.* New York: Holt, Reinhart, & Winston Publishing.

Hammond, K. R. (2001). Expansion of Egon Brunswik's Psychology, 1955-1995. In K. R. Hammond & T. R. Stewart (Eds.), *The Essential Brunswik* (pp. 465-478).

Hammond, K. R., & Summers, D. A. (1972). Cognitive control. *Psychological Review, 79* (1), 58-67.

Hammond, K. R. (1996). *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*. New York: Oxford University Press.

Hammond, K. R. (2001). How probabilistic functionalism advances evolutionary psychology. In K. R. Hammond & T. R. Stewart (Eds.), *The Essential Brunswik* (pp. 434-441).

Hammond, K. R., Hamm, R. M., Grassia, J., & Pearson, T. (1987). Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. *IEEU Transactions On Systems, Man, & Cybernetics, 17* (5), 753-770.

Hammond, K. R., Stewart, T. R., Brehmer, B., & Steinmann, D. O. (1975). Social judgment theory. In M. Kaplan & S. Schwartz (Eds.), *Human judgment and decision processes*. New York. Academic Press (pp. 271-312).

Hammond, K. R. (1988). Judgment and decision making in dynamic tasks. *Information and Decision Technologies, 14*, 3-14.

Hammond, K. R. (1990). Intuitive and analytical cognition: Information models. In A. P. Sage (Ed.), *Concise encyclopedia of information processing in systems and organizations*. Oxford: Pergamon Press, 306-312.

Hoffman, P. J., Slovic, P., & Rorer, L. G. (1968). An analysis of variance model for the assessment of configural cue utilization in clinical judgment. *Psychological Bulletin, 69*, 338-349.

Holzworth, R. J. (2001a). Judgment analysis. In K. R. Hammond & T. R. Stewart (Eds.), *The Essential Brunswik* (pp. 324-327).

Holzworth, R. J. (2001b). Multiple cue probability learning. In K. R. Hammond & T. R. Stewart (Eds.), *The Essential Brunswik* (pp. 348-350).

Juslin, P., Winman, A., & Olsson, H. (2000). Naïve empiricism and a priori dogmatism in confidence research: A critical examination of the hard-easy effect. *Psychological Review, 107*, 384-396.

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions.

Cognition, 11, 123-141.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 26* (4), 349-370.

Libby, R. (1976). Man versus model of man: Some conflicting evidence. *Behavior and Human Performance, 16*, 1-12.

Lichtenstein, S., & Fischoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance, 20*, 159-183.

Lindell, M. K. (1976). Cognitive and outcome feedback in multiple-cue probability learning tasks. *Journal of Experimental Psychology: Human Learning and Memory, 2*(6), 739-745.

Maki, R. H., Foley, J. M., Kajer, W. K., & Thompson, R. C. (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory & Cognition, 16* (4), 609-616.

Manning, B. H., & Payne, B. D. (1996). *Self-talk for teachers and students: Metacognitive strategies for personal and classroom use.* Boston: Allyn & Bacon Press.

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory & Cognition, 19*, 851-861.

Morris, C. C. (1990). Retrieval processes underling confidence in comprehension judgments. *Journal of Experimental Psychology: Learning, Memory, and*

Cognition, 16 (2), 223-232.

Naylor, J. C., & Clark, R. D. (1968). Intuitive inference strategies in interval learning tasks as a function of validity magnitude and sign. *Organizational Behavior and Human Decision Processes, 3*(4), 378-399.

Naylor, J. C., & Schenck, E. A. (1968). The influence of cue redundancy upon the human inference process for tasks of varying degrees of predictability. *Organizational Behavior and Human Decision Processes, 3*(1), 47-61.

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL" effect. *Psychological Science, 2* (4), 267-270.

Peters, J. T., Hammond, K. R., & Summers, D. A. (1974). A note on intuitive vs analytic thinking. *Organizational Behavior and Human Decision Processes, 12*(1), 125-131.

Pressley, M., & Ghatala, E. S. (1990). Self-regulated learning: Monitoring learning from text. *Educational Psychologist, 25*, 19-33.

Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory & Cognition, 18*, 435-451.

Reilly, B. A., & Doherty, M. E. (1989). A note on the assessment of self-insight in judgment research. *Organizational Behavior and Human Decision Processes, 44*, 123-131.

Reilly, B. A., & Doherty, M. E. (1992). The assessment of self-insight in judgment policies. *Organizational Behavior and Human Decision Processes, 53*, 285-309.

Reilly, B. A. (1996). Self-insight, other-insight, and their relation to interpersonal conflict. *Thinking & Reasoning, 2*, 213-223.

Sawyer, J. E. (1991). Hypothesis sampling, construction, or adjustment: How are inferences about nonlinear monotonic contingencies developed? *Organizational Behavior and Human Decision Processes, 49*, 124-150.

Schaefer, P. S., Williams, C. C., Goodie, A. S., & Campbell, K. (2002). Metacognition and the Five Factor Model: Predicting Overconfidence. Submitted.

Schenck, E. A. (1969). A study of the independent effect of cue redundancy upon the human inference process in tasks of varying predictability. *Dissertation Abstracts International, 29*(9-B), 3516.

Schmitt, N., & Dudycha, A. L. (1975). A reevaluation of the effect of cue redundancy in multiple-cue probability learning. *Journal of Experimental Psychology: Human Learning and Memory, 1*(3), 307-315.

Schmitt, M. C., & Baumann, J. F. (1986). How to incorporate comprehension monitoring strategies into basal reader instruction. *Reading Teacher, 40* (1), 28-31.

Schraw, G., & Roedel, T. R. (1994). Test difficulty and judgment bias. *Memory & Cognition, 22*, 63-69.

Schraw, G., Potenza, M., & Nebelsick-Gullet, L. (1993). Constraints on the calibration of performance. *Contemporary Educational Psychology, 22*, 63-69.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19* (4), 460-475.

Schraw, G. (1997). The effect of generalized metacognitive knowledge on test performance and confidence judgments. *The Journal of Experimental Education, 65* (2), 135-146.

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science, 26*, 113-125.

Schraw, G., & Nietfield, J. (1998). A further test of the General Monitoring Skill hypothesis. *Journal of Educational Psychology, 90* (2), 236-248.

Schraw, G., Dunkle, M. E., Bendixen, L. D., & Roedel, T. D. (1995). Does a general monitoring skill exist? *Journal of Educational Psychology, 87*, 433-444.

Sheets, C. A., & Miller, M. J. (1974). The effect of cue-criterion function form on multiple-cue probability learning. *American Journal of Psychology, 87*(4), 629-641.

Smedslund, J. (1955). *Multiple probability learning*. Oslo, Norway: Oslo University Press.

Sniezek, J. (1986). The role of variable labels in cue probability learning tasks. *Organizational Behavior and Human Decision Processes, 38*, 141-161.

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

Steinmann, D. O. (1974). Transfer of lens model training. *Organizational Behavior and Human Decision Processes, 12*(1), 1-16.

Stewart, T. R. (2001). The lens model equation. In K. R. Hammond & T. R. Stewart (Eds.), *The Essential Brunswik* (pp. 357-362).

Stimson, M. J. (1998). Learning from hypertext depends on metacognition.

Unpublished doctoral dissertation. University of Georgia.

Suantak, L., Bolger, F., & Ferrell, W. R. (1996). The hard-easy effect in subjective probability calibration. *Organizational Behavior and Human Decision Processes, 67* (2), 201-221.

Summers, S. A., Summers, R. C., & Karkau, V. T. (1969). Judgment based on different functional relationships between interacting cues and a criterion. *American Journal of Psychology, 82*, 203-211.

Swanson, H. L. (1990). Influence of metacognitive knowledge and aptitude on problem solving. *Journal of Educational Psychology, 82* (2), 306-314.

Uhl, C. N. (1963). Learning of interval concepts: I. Effects of differences in stimulus weights. *Journal of Experimental Psychology, 66*(3), 264-273.

Veenman, M. V. J., Elshout, J. J., & Meijer, J. (1997). The generality vs. domain specificity of metacognitive skills in novice learning across domains. *Learning & Instruction, 7*, 187-209.

Von Schantz, T., Gorannson, G., Anderson, G., Froberg, I., Grahn, M., Helgee, A., & Witzell, H. (1989). Female choice selects for a viability male trait in pheasants. *Nature, 337*, 166-169.

Wiggins, N., & Hoffman, P. J. (1968). Three models of clinical judgment. *Journal of Abnormal Psychology, 73*, 70-77.

Wigton, R. S., Patil, K. D., & Hoellerich, V. L. (1986). The effect of feedback in learning clinical diagnosis. *Journal of Medical Education, 61* (10), 816-22.

APPENDIX A

DEFINITIONS

Achievement: the degree of correlation between a judge's responses to sets of cues and the observed criterion values that accompany those sets of cues.

Cognitive control: a measure of the similarity between a judge's responses in a judgment task and the predictions of those responses made by a specific model. It is measured by the multiple correlation between judgments and the predictions of those judgments by the judge's policy equation. Higher levels of cognitive control indicate more consistent application of an individual's *policy*.

Criterion: this term is used to refer to the value or state within the ecology which the judge is attempting to achieve when making a judgment.

Cues: any numerical, graphical, verbal, pictorial, or other sensory information which is available to a judge for potential use in forming a judgment for a specific case and/or which is available in the ecology for making predictions about some criterion.

Cue redundancy: the extent to which cues covary. Usually expressed as the average intercorrelation of the cues.

Ecological regression model: term used to refer to the model that represents the relationship between cues (as the predictor variables) and the criterion (as the predicted variable).

Ecological validity: correlation between the values a particular cue takes on and the values a criterion takes on.

Ecology: the totality of cues, criterion, and all of their interrelationships.

Functional validity: the correlation between the values a particular cue takes on and the values judgments take on.

Function form: the mathematical relationship between the values of a cue and the values of judgments. Function forms may be linear or nonlinear. Function forms are usually graphically depicted.

Judgments: an explicit indication of a judge's appraisal of a set of cue values with respect to some dimension of interest (the criterion).

Matching: a concept used to refer to the extent to which predictions from the judge's captured policy are correctly related to the predictions made by the model of the ecology. It is defined by the correlation between the predicted values of policy (predicted judgments) and ecology (predicted criterion values).

Predictability: a measure of the similarity between the observed criterion values in an ecology and the prediction of those values made by a specific model. It is measured by the multiple correlation between observed criterion values and predictions of those values by the regression model of the ecology.

Probabilistic Functionalism: a term used to refer to Brunswik's concern with the probabilistic (as opposed to deterministic) view of the ecology within which most organisms operate. The term functionalist indicates his stress upon the utilitarian, adjustment centered idea of behavior.

Policy: a term used to indicate the model used to represent the judgment process of a judge. Typically this refers to the multiple regression equation fitted to a person's judgments. The cues are entered as predictor variables and the judgments as the predicted variable.

Subjective weights: relative weights derived from importance ratings assigned to cues by participants. Correspondence between subjective weights and regression weights is one measure of insight.

APPENDIX B

PROBABILISTIC FUNCTIONALISM

Probabilistic functionalism examines functioning in an uncertain environment in a variety of ways. One way is to draw a distinction between true states of nature and information about those true states that reach our sensory organs. These true states of nature are referred to as *criteria*, and the information that reach our sensory organs are known as *cues*. The extent to which a given cue is predictive of the criterion is known as the cue's *ecological validity*. Conversely, the extent to which a series of judgments is based upon a cue is known as the *functional validity* of the cue.

Any given cue is always a less-than-perfect representation of a criterion. The relationship between cue and criteria are, rather, of a probabilistic nature. Some cues are more predictive than other cues. Adaptive behavior is defined as the extent to which the ecological validities and the functional validities of a cue correspond to each other.

Another important feature of probabilistic functionalism is the fact that, although early psychological research on judgment utilized cues that were constrained to be orthogonal, cues in everyday decisions often exhibit *cue redundancy* (Brunswik, 1939; Brunswik, 1952).

APPENDIX C

MCPL RESEARCH

Feedback

MCPL research indicates that outcome feedback alone is of minimal efficacy, restricted to fairly simple tasks involving cue-criterion relationships that are positive and linear in form (Lidnell, 1976; Steinmann, 1974; Wigton & Hoellerich, 1984). Nonlinear cue-criterion relationships require many more trials if outcome feedback alone is provided (Wigton, Patil, & Hoellerich, 1986).

Cue-Criterion Relationships

Although individuals can learn to utilize cues that are related to a criterion in a nonlinear fashion (Brehmer, 1969), individuals learn the relationship between cues and the criterion more easily if it is linear in nature (Brehmer, 1976; Brehmer & Svensson, 1976; Naylor & Clark, 1968; Bjorkmann, 1965). The most commonly used nonlinear relationship has been a quadratic one (de Klerk & Oppe, 1972; de Klerk & Vroon, 1974; Cooksey, 1996).

More specifically, it appears that individuals find it easiest to learn a task that is positive and linear in nature. Negative linear tasks are more difficult, followed by nonlinear cue usage (Brehmer, 1971; Eisler & Spolander, 1970). When learning a quadratic function, individuals learn a U-shaped function more readily than an inverted U-shaped function (Brehmer, 1974a; Sniezek, 1986; Sniezek & Naylor, 1978; Sawyer, 1991). In fact, it appears that even when individuals are aware that the relationship between the criterion values and cue values is best expressed as a quadratic one, individuals focus initially on learning the linear portion (Earle, 1973; Sheets & Miller, 1974; Summers, Summers, & Karkau, 1969).

Cue Redundancy and Predictability

Another factor that appears to impact performance is *cue redundancy*, or the degree to which cues are intercorrelated. In general, the greater the degree of intercorrelation among the cues, the greater the performance (Naylor & Schenck, 1968; Schenck, 1969). However, it should be noted that the relationship between cue intercorrelation and performance is dependent upon the factor of predictability.

Predictability may be thought of as the amount of variance among criterion values that are predicted by variance among cue values (the R^2 of the regression model expressing the task ecology). In a highly predictable task, greater performance is seen when cue redundancy is low than when it is high (Uhl, 1963; Schmitt & Dudycha, 1975).

APPENDIX D

LENS MODEL EQUATION

Cognitive Control and Predictability

The parallel between cognitive control and predictability helps to underscore the symmetrical emphasis placed upon the ecology and the organism in the LME. Cognitive control may be thought of as analogous to predictability in that, in both instances, the correlation between a set of actual and predicted values expresses both measures. However, the important distinction is that while predictability is a descriptive statistic of an ecology, cognitive control arises out of a mental process.

Therefore, while predictability is invariant for a given experimental condition, there are as many measures of cognitive control for an experimental conditions as there are participants in that condition. Cognitive control is so named because it reflects the extent to which an individual consistently combines cue information in the same manner. If, for example, an individual integrated the cues with perfect cognitive control across all trials, there would be a 1.00 correlation between the predicted judgments and the actual judgments.

Achievement versus Matching

Cognitive control variance drives the difference between achievement and matching values. The familiar distinction between noise and signal is illustrative. The difference in the way that cues are related to observed criterion values and predicted criterion values revolves around the amount of noise, or unexplained variability, present in both relationships. The regression equation removes any noise from the situation and, through perfectly consistent application, generates a set of values in such a way that the relationship between the cues and predicted criterion values is a relationship of pure signal.

This situation is mirrored in the participant part of the lens model as well. In many situations, especially in the current circumstance where predictability in all of the ecologies was nearly perfect, the major source of noise is low cognitive control. That is, participants usually have some amount of signal (explained variance) present, but there is also a large amount of noise (unexplained variance) due to inconsistent application of a strategy. A higher cognitive control value indicates a greater amount of signal, and a lesser amount of noise.

Because matching is the correlation between the predicted criterion values of the ecology and the predicted judgments of an organism, matching is an indication of how correct the participant's policy would be, *if* the environment were perfectly predictable, and *if* the participant's policy were enacted with perfect consistency (i.e., cognitive control of 1.00).

Interdependence of LME Indices

It should be noted that these performance indices are not totally independent. For example, there is one instance in which, of necessity, the matching and achievement indices will be precisely the same. This occurs when achievement is 1.00; that is, when an individual's judgments are exactly the same as the observed criterion values.

In such a circumstance, the ecological regression model and the participant's policy will be exactly the same. The predictor values are the same in any case, because both the participant's judgments and observed criterion values are regressed upon the same cue values. When the judgments and observed criterion values are the same, and both are being regressed upon the same cue values, then naturally the predicted judgments and predicted criterion values will also be the same. Thus, when achievement is 1.00, so too is matching.

However, because humans are often inconsistent in their application of a policy (i.e., have low cognitive control), there is almost always a discrepancy between matching and achievement. In fact, a common finding in the literature is that utilizing the policy of a judge rather than the judge her or himself to make subsequent decisions often results in greater accuracy than utilizing the judge him or herself because of the removal of human inconsistency (Bowman, 1963; Cooke, 1967; Libby, 1976; Camerer, 1981b). A couple of examples may help clarify this distinction between matching and achievement.

It is perfectly plausible, for example, for participant A to have a matching index that is higher than participant B, but for participant B to have an achievement index that is higher than participant A. This could arise because the policy of participant A was closer to that of the ecological regression model than the policy of participant B, but participant B enacted his policy with greater cognitive control. In more common language one could say that when participant A combined cues in a systematic fashion, he did so more accurately than when participant B combined cues systematically. However, the majority of the time participant A combined cues in a nonsystematic way. Therefore, having the correct policy alone is not enough: one must also apply it consistently to attain higher levels of achievement.

Conversely, it is also possible that low cognitive control can be advantageous if a wildly incorrect policy is being used. For example, participants A and B might be utilizing cues in a positive linear fashion when the cues are in fact inversely related to some criterion of interest. Assuming that participant A is still enacting his policy with less cognitive control than participant B, what will happen?

In such a situation, it is possible that the reverse situation emerges. That is, the participant who is enacting a policy with less cognitive control might exhibit higher achievement than the participant enacting a policy with greater cognitive control. After all, if a policy is wildly incorrect, perfect cognitive control in its application will result in wildly inaccurate judgments. Both of these examples serve to illustrate the compensatory nature of Brunswikian decision making, in which high values on one performance index can partially compensate for low values on another.

APPENDIX E

INSIGHT

Common Measures of Insight

Insight has been measured in diverse ways. Some studies have utilized think-aloud protocols in which experts described their policy by declaring, during the judgment process, what cues were important and how they were utilizing them (Ericsson & Simon, 1993; Einhorn, Kleinmuntz, & Kleinmuntz, 1979; Armelius & Armelius, 1975a; Armelius & Armelius, 1975b). In some cases, individuals are shown a matrix of linear regressions expressing judgment policies and are asked to indicate which one is their own (Reilly & Doherty, 1992).

However, the most common way to measure insight recently has been to require individuals to rate, on a likert scale, how important they think that various cues were when integrating and predicting criterion values (Gray, 1979; Brehmer, 1977; Reilly & Doherty, 1989; Reilly, 1996). These weights are known as *subjective weights*. Insight is often expressed as the similarity (sometimes a correlation) between the subjective weights derived from the subject and the statistically captured regression weights from their policy.

Limitations of Insight in Likert Scale Subjective Weights

However, the process by which these subjective weights are normally rendered comparable to standardized beta weights is only appropriate when the task ecology is linear (Hammond, Stewart, Brehmer, & Steinmann, 1975; see also Cooksey, 1996, pp.

177, for a discussion of how weight and function forms interact with each other) As noted by other researchers (Haarbauer, 1996), this is especially troublesome for CCT. CCT posits that nonlinear cue-criterion function forms should result in greater levels of insight, because nonlinear cue usage is associated with more analytical tasks (see Tables 1 and 2). However, it is quite possible that insight per se would not be affected by utilizing the standard method of comparing subjective weights to the standardized beta weights, but that violation of the linearity assumption would result in greater disparity between the two sets of weights.

Hence, the most common way of measuring insight was inappropriate in the current experiment, wherein one half of all conditions were nonlinear (quadratic) in nature. Therefore, another method of examining insight was utilized that allowed for the assessment of subjective weights in nonlinear environments.

Current Measurement of Insight

The traditional method of comparing subjective weights to statistically captured weights was inappropriate for the current experiment However, the judgment analysis software package utilized in the proposed experiment (POLICY PC) allows for specification of the cue-judgment function forms as well as assignation of relative weights.

While this avoided the problem of confounding the cue weights with the cue-criterion (or in this case cue-judgment) function forms, it raised another issue of concern. For participants to correctly utilize this method would require specification of the mean and standard deviation of the last set of judgments, as well as correct specification of the judgment-criterion function form.

This would place more of a burden upon these subjects than participants in previous experiments that utilized the likert scale method of eliciting subjective cue weights. Therefore, the researcher utilized Policy PC to specify the cue-judgment function form that best expressed the participant's policy, specified the correct indicators of central tendency for the participants' judgments, and entered the participant elicited relative weights. Policy PC then generated a set of predicted judgments from that process, which were then correlated with the participant's predicted judgments.

This rationale is consistent with that underlying the computation of the matching index of performance. Both involve constructing a metric for expressing the similarities of rules. Both also take into account the inconsistency with which such a rule might be applied by fallible human subjects (Camerer, 1981b).

APPENDIX F

LINEAR MODELS OF JUDGMENT

Generality of Linear Models of Judgment

Linear models describe many judgment activities quite well, whether one is measuring group communication among apes (Byrne, 1995) or pheasant mate selection (Von Schantz, Gorannson, Anderson, Froberg, Grahn, Helgee, & Witzell, 1989). Hammond (2001) has argued vigorously that there is ample evidence across species for processing in a linear, equal weights fashion.

Robustness of Linear Models

There are also several possible advantages to the use of such a model. The first is that of cognitive economy: an equal weighting model would presumably take up fewer resources than one that requires disparate weighting (Camerer, 1981b). In addition, when cross-validation occurs, equal weights models are not as subject to shrinkage in R^2 as are regression models with disparate cue weightings. Thus, equal weights models would have an advantage as a cross-domain judgment policy.

Further, Hammond (1996) has noted that humans often create such models (linear, multiple fallible cues probabilistically related to a criterion) to make judgments and express variation in the environment. Reilly & Doherty (1992) have noted that the finding that individuals can often recognize their judgment policies when described by a linear equation is in itself a roundabout argument that their decision processes were sufficiently captured by a linear regression model (Wiggins & Hoffman, 1968; Camerer, 1981b; Camerer & Johnson, 1997; Hoffman, Slovic & Rorer, 1968).

This evidence is consistent with Hammond's contention that because humans hardwired to process things in a linear, equal-weights fashion, it is possible to do so with little or no insight as to one's actual judgment policy (Hammond, 1996). That is, decision making in such an environment is something akin to an automatic process.

However, encountering situations in which the cue-criterion function form is nonlinear requires more of a controlled process, and thus also requires a higher level of insight for consistent application. This is often phrased in terms of difficulty or complexity: the more difficult (i.e., nonlinear) a task is, the greater the amount of conscious control that is necessary for better performance.