An evaluation between confidence judgments and differences in monitoring ability

Brett Douglas Campbell
University of Nevada, Las Vegas

Follow this and additional works at: https://digitalscholarship.unlv.edu/rtds

Repository Citation
https://digitalscholarship.unlv.edu/rtds/1352

This Thesis is brought to you for free and open access by Digital Scholarship@UNLV. It has been accepted for inclusion in UNLV Retrospective Theses & Dissertations by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
AN EVALUATION BETWEEN CONFIDENCE JUDGMENTS AND DIFFERENCES IN MONITORING ABILITY

by

Brett Douglas Campbell

Bachelor of Arts
Weber State University
1990

A thesis submitted in partial fulfillment of the requirements for the

Master of Arts Degree
Department of Educational Psychology
College of Education

Graduate College
University of Nevada, Las Vegas
August 2002
Thesis Approval
The Graduate College
University of Nevada, Las Vegas

The Thesis prepared by

Brett Campbell

Entitled

An Evaluation Between Confidence Judgments and Differences in Monitoring Ability

is approved in partial fulfillment of the requirements for the degree of

Master of Science in Educational Psychology

Examination Committee Chair

Dean of the Graduate College

Examination Committee Member

Examination Committee Member

Graduate College Faculty Representative
ABSTRACT

An Evaluation Between Confidence Judgments and Differences in Monitoring Ability

by

Brett Douglas Campbell

Dr. Alice Corkill
Associate Professor of Educational Psychology
University of Nevada, Las Vegas

Metacognition is the awareness of a person's thinking and the ability to regulate that thinking. This study examines how confident a person is at monitoring progress during two cognitive tasks. The first task was the feeling-of-knowing paradigm, when participants judge how accurately they can retrieve an answer in a recognition task, when they cannot retrieve the answer in free recall. The second task was a paired-associate interference paradigm. Participants were presented with a list of noun-noun pairs twice and then a new list was presented for recall. Participants were placed into three groups (high, medium, low) based on responses to the Metacognitive Awareness Inventory. Results showed no significant differences between groups in either task. The results are discussed in terms of measuring strategy knowledge instead of monitoring.
# TABLE OF CONTENTS

ABSTRACT .................................................................................................................... iii

LIST OF TABLES ........................................................................................................... vi

LIST OF FIGURES ...................................................................................................... vii

CHAPTER 1 INTRODUCTION ..................................................................................... 1
  Metacognitive Knowledge ...................................................................................... 2
  Metacognitive Control ............................................................................................ 10
  Monitoring Ability ................................................................................................. 14
  Feeling-of-Knowing ............................................................................................... 16
  Target Availability ................................................................................................. 19
  Cue Familiarity ....................................................................................................... 23
  The Present Study .................................................................................................. 25
  Metacognitive Awareness Inventory .................................................................... 27
  Need for Cognition ................................................................................................. 30

CHAPTER 2 METHOD ................................................................................................. 35
  Subjects .................................................................................................................... 35
  Materials .................................................................................................................. 35
  Metacognitive Awareness Inventory ..................................................................... 35
  Need for Cognition Scale ....................................................................................... 36
  Recall ....................................................................................................................... 36
  Recognition ............................................................................................................. 37
  Paired-associate ..................................................................................................... 37
  Procedure ................................................................................................................. 38

CHAPTER 3 RESULTS ................................................................................................. 41

CHAPTER 4 DISCUSSION ........................................................................................... 47
  Limitations .............................................................................................................. 51
  Future Directions ................................................................................................... 54

APPENDIX I INVENTORIES AND TESTS ................................................................. 59

APPENDIX II TABLES AND FIGURES REFERENCED IN TEXT ............................ 78

REFERENCES ............................................................................................................. 91

iv

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
LIST OF TABLES

Table 1  Mean and standard deviations of groups on MAI Factor 2  ......................... 79
Table 2 Bias means on feeling-of-knowing task ......................................................... 80
Table 3 Bias means on paired-associates task ............................................................ 81
Table 4 Calibration means on feeling-of-knowing task ............................................. 82
Table 5 Calibration means on paired-associates task ............................................... 83
Table 6 Means and standard deviations of groups on MAI Factor 1 ......................... 84
Table 7 Means and standard deviations of groups on total MAI score ...................... 85
Table 8 Correlation matrix between MAI and NFC .................................................. 86
LIST OF FIGURES

Figure 1 Bias Means by Ability Group ................................................................. 87
Figure 2 Calibration Means by Ability Group ..................................................... 88
Figure 3 Paired-associates Bias Means ............................................................... 89
Figure 4 Paired-associates Calibration Means ................................................... 90
CHAPTER 1

INTRODUCTION

Metacognition has been defined as knowledge and control over personal cognitive faculties and content (Brown, 1987; Flavell, 1987; Flavell & Wellman, 1977; Swanson, 1990). Flavell (1987) described metacognition as a model of cognitive monitoring, denoting the cognitive and affective experiences involved. He further indicated that self-analysis leads to planning and learning from such reflective experiences.

The first comprehensive theory of metacognition was published in 1979 (Flavell). The theory was based on previous research in cognition. Metacognition research has become one of the most popular research areas involving child development, education, and counseling. It has been related to reading comprehension (e.g. Brown, 1975, 1978), problem solving (e.g. Brown, 1987; Cornoldi, 1998; Davidson, Deuser & Sternberg, 1994), and storage and retrieval tasks (e.g. Nelson, 1984; Pressley, Borkowski & Sullivan, 1985).

Flavell compared metacognition to Piaget's reflective abstraction. As an individual becomes increasingly aware of their cognitive abilities, they gauge their capacity for storing information in retrieval and encoding tasks. This understanding of personal abilities and limitations occurs in both formal teaching environments and through self-
reflection. Piaget defined reflective abstraction as "an internal mechanism of reflection on its reflection" (as cited in Montangero & Maurice-Naville, 1997).

Metacognition has been divided into two components: 1) the knowledge of personal cognitive abilities and processes and 2) the ability to control and modify these cognitive activities (Brown, 1987; Corkill & Koshida, 1993; Flavell, 1987; Jacobs & Paris, 1987; Schraw & Moshman, 1995). Knowledge, or awareness, about cognitive abilities and processes has been labeled metacognitive knowledge. Controlling or regulating cognitive abilities and processes have been identified as metacognitive control.

The remainder of this chapter begins with a review of metacognitive knowledge. Research in this area has focused on the relationship between intellectual aptitude and metacognition and metacognitive strategy knowledge. The next section considers metacognitive control, investigating how metacognitive control has been measured. A more in-depth look at research associated with on-line monitoring as an aspect of metacognitive control is included. One approach to measuring metacognitive monitoring has been the examination of feeling-of-knowing judgments. Following is a review of the two current theories explaining feeling-of-knowing judgments. The justifications for the present study are then elaborated. Finally, there is an examination of two inventories used in this study—the Metacognitive Awareness Inventory and the Need for Cognition scale.

Metacognitive Knowledge

Flavell's metacognitive theory is examined first. A second theory of how metacognitive knowledge has been defined is then presented. Research into metacognitive knowledge can be grouped into two areas. The first is the relationship between aptitude
and metacognitive knowledge. A review of the work of Swanson and others examines this relationship. The second area is the examination of strategy knowledge.

Flavell first identified metacognitive knowledge after studying metamemory in children (Flavell, 1979; 1987). The initial definition of metacognitive knowledge suggested it was comprised of three components: person, task, and strategy variables (Flavell, 1979; 1987; Flavell & Wellman, 1977). Person variables identify the awareness people have about the contents of their memory. It also includes knowledge about how to best encode material to be learned and how to retrieve that material. This includes comparing one’s abilities to the abilities of others. Tasks variables are defined as the knowledge one has regarding differences between specific tasks, as well as the effects of storing and retrieving information according to task difficulty. Task variables include the ability to find relationships between sets of information or comparing the current task with previous mental experiences. Strategy variables include the knowledge one has of specific strategies learned via experience or formal training and how each strategy may apply to different tasks.

Others have avoided Flavell’s nomenclature, focusing on individual awareness instead. Metacognitive awareness is divided into declarative knowledge, procedural knowledge, and conditional knowledge (Chi, 1987; Gavelek & Raphael, 1982; Jacobs and Paris, 1987; Schraw & Dennison, 1994; Schraw & Graham, 1997). Declarative knowledge refers to knowledge about self and what specific factors affect performance, such as capacity, use of rehearsal, and availability of different strategies. Procedural knowledge refers to use and selection of strategies. Conditional knowledge refers to when and why an
individual would use declarative and procedural knowledge, including which strategies are effective, when they are effective, and when they are appropriate.

Swanson (1987; see also Daneman & Carpenter, 1980) compared metacognitive knowledge variables with intelligence in memory recall. Questionnaires, based on the work of Flavell measuring metacognitive knowledge, were given to children of normal and high intelligence. After the questionnaires were administered, two recall tasks were completed. In the first task, children were asked to organize a series of 15 pictures in any order that would help them recall the greatest number of pictures. The second task required the children to listen to a story and then recall as many of the story details as possible. Swanson found Flavell's strategy and task variables were better predictors of recall than the person variable or intelligence—as determined by standardized verbal ability tests.

Additional research has demonstrated that metacognitive ability can compensate for cognitive aptitude (Swanson, 1990). Children were placed into one of four groups: 1) high intelligence, high monitoring ability; 2) high intelligence, low monitoring ability; 3) low intelligence, high monitoring ability; and 4) low intelligence, low monitoring ability. Each child then had to perform two tasks. In the first, the subject had to deduce which two colorless liquids would create a colored liquid from four options. In the second task, the child had to control the speed of a pendulum by varying the weight at the end of a string and by varying the length of the string. The two high metacognitive groups did better at both tasks than either low metacognitive group. Of the two low metacognitive ability groups, the high intelligence subjects performed better than those with low intelligence.
Intellectual aptitude was considered less vital than metacognitive ability in these reasoning tasks.

Swanson, Christie, and Rubadeau (1993) compared children who were average readers with children who were learning disabled. While there were differences in reading comprehension, vocabulary, and reading rates, they performed similarly in metacognitive activities. The similarity in performance scores between average readers and those who were classified as learning disabled suggested that children with learning disabilities can and do use metacognitive skills to compensate for specific domain deficiencies.

Swanson and Trahan (1996) compared metacognitive differences in solving analogies in children between 12 and 14 years of age. The groups compared were average achievers, gifted, learning disabled, and mentally retarded. The gifted children had the highest scores of all groups in both the metacognition inventory and solving analogies. Those with mental retardation performed significantly lower in metacognition scores and analogies. The normal and learning disabled groups had similar scores on the analogies task. Individuals with learning disabilities were able to compensate with metacognitive strategies; whereas, mentally retarded individuals apparently could not employ compensatory strategies. Flavell’s strategy knowledge, as measured by Swanson’s inventory, was significantly correlated to success in the analogy task.

In another study, Keeler and Swanson (2001) examined ten-year old children with respect to metacognitive knowledge about math. Additional variables of interest included spatial and verbal ability. Subjects completed two tasks. The first, a digit sentence span task, required subjects to read sentences that contained several digits. Subjects were then asked to recall all digits from the sentences. In the second task, subjects were asked to
study a path on a map that had no labels. The recall test required subjects to draw the path on a blank map. In each task subjects were given four strategies from which to choose: rehearsal, clustering, association, and elaboration. Interviewing the child during the task assessed which strategy was being used. The results suggested that only strategy knowledge and working memory span accounted for significant proportions of variance in successful completion of both tasks. Those children who stayed with one strategy during the task had better performance scores. However, the strategy chosen was dependent on an individual’s working memory span.

The utility of strategy instruction, as a metacognitive component, has been effectively demonstrated (e.g. Baker & Brown, 1984; Brown, 1975; 1987; Pressley & Ghatala, 1990; Pressley, 1995; Slife, Weiss & Bell, 1985). Strategy instruction begins to become effective between second and third grade (Brown, Campione, & Murphy, 1977; Devolder, Brigham & Pressley, 1990; Pressley, 1994; Schneider & Pressley, 1997). For example, Brown and colleagues trained children between the mental ages of 6 and 8 to use mnemonic strategies. The educable children with a mental age of 6 could only remember the strategy at immediate testing. Those with the mental age of 8 could remember the strategy when tested one year later.

Paris and Jacobs (1984) taught third and fifth grade children metacognitive strategies in reading comprehension. Half-hour lessons were conducted twice a week for the entire class for four months. Reading comprehension was measured through four tasks, a structured interview with the investigator, the Gates-MacGinitie reading comprehension test, a cloze task, and an error detection test. The cloze task consisted of a passage with several words missing. The children then filled in the correct word. This task was
considered a measure of literal comprehension, inferential comprehension, identifying the main idea, and structural awareness. The error detection test consisted of a reading passage with several inconsistencies within the passage. Subjects were to underline those words or sentences that did not make sense. Children were tested in the fall and again in the spring of the school year. Children who received strategy instruction made the most improvement in reading comprehension regardless of age. Participants were divided into groups of high, middle, and low awareness according to their responses to the interview. Paris and Jacobs found the high awareness group scored the highest on the Gates-MacGintie, the cloze task, and the error detection test.

A follow-up study (Jacobs & Paris, 1987) examined comprehension strategies in third and fifth grade students. Instead of interviewing the children, they created an inventory. Classes were assigned to either receive strategy instruction or no strategy instruction. The instruction included explicit teaching of comprehension strategies, discussions, and practice sessions. Sixty lessons were taught over the course of a school year. The Gates-MacGintie standardized test was used again to ascertain reading awareness. Children were tested towards the end of the school year and the results were compared to pre-tests scores. This research found that subjects who received comprehension strategy instruction had greater reading awareness in both grades. As would be expected, the fifth grade children had higher scores than the third grade children. In addition, the third grade girls had higher comprehension scores than the boys.

Delclos and Harrington (1991) gave fifth and sixth grade children strategy instruction in solving logic problems. Students had to solve problems in a computer game format. Each child was placed into one of three groups: problem-solving, monitored
problem solving, or a control group. The two groups receiving strategy instruction received two hours of instruction. The monitored group also received a booklet consisting of questions designed to reinforce monitoring strategies during the practice phase. All three groups participated in a practice phase. Both instruction groups both outperformed the control group. The group receiving strategy instruction with monitoring did significantly better than the students receiving strategy instruction alone.

Unfortunately, children are not inclined to employ successful strategies until late childhood or late adolescence (Baker, 1985; Brown & Smiley, 1978; Schneider, 1998). Brown and Smiley (1978), for example, tested students in grade five, grades seven and eight, grades eleven and twelve, and college students in reading comprehension and use of strategies. Participants were presented with a story and then later tested on recalling the theme of the story. Grade eleven and twelve students and the college students were the most efficient at identifying themes. The seventh and eighth grade students were more proficient than fifth graders at the recall task, but did not do well as the oldest participants. When subjects were divided into spontaneous strategy or induced strategy (told to underline key concepts in a reading passage) and told to identify important aspects in a passage, the spontaneous strategy users were more proficient in identifying the key themes, while the induced strategy users appeared to select themes in a random fashion.

McGivern and colleagues (McGivern, Levin, Pressley, & Ghatala, 1990) divided second graders, seventh graders, and college students into three learning groups: 1) the first group was explicitly taught to self-monitor, 2) the second group was given the monitoring instruction plus an opportunity to practice on-line monitoring, and 3) the third group was given no instruction. All three groups watched a video of a model learning
noun pairs. The model demonstrated two strategies, repetition or elaboration. In the repetition strategy, subjects were shown the model repeating paired words three times. In the elaboration strategy, the model produced a sentence incorporating each word pair. The investigators interviewed the participants about each strategy. Subjects were then tested twice in a paired-associate paradigm. In one test, they were instructed to use the repetition strategy and in the second test they were instructed to use the elaboration strategy. Afterwards each student was asked which strategy was more useful. The college students suggested the elaboration strategy was more efficient than the repetition strategy. The seventh grade students who received instruction performed as well as the college students in identifying elaboration as a better strategy. The seventh grade with no instruction group believed both strategies to be equally efficient. The second grade students showed poor strategy monitoring overall. This suggests that strategy instruction may be a necessary component in facilitating between two strategies in children as late as twelve and thirteen years old.

Son and Metcalfe (2000) examined study-time-allocation strategies. Individuals were tested using biographical essays and Japanese haiku. In the biography test subjects read eight one-paragraph biographies, then judged each biography according to how easy it would be to learn and how interesting each paragraph was. For those given poetry, subjects read 49 haikus and made ease-of-learning judgments and judgments-of-interest. An ease-of-learning judgment is a rating of how easy the subject believes the information is to remember or to learn. The judgment-of-interest asks subjects to indicate level of interest in a particular reading passage or poem. Participants were placed in either a free-reading condition or in a study-for-test condition. All groups were tested for recall. Students who
knew they would be tested focused on the material that was judged easier to learn. The results suggested that focusing on easier items reinforced learned items, guaranteeing a higher test score.

In review, recent studies have supported Flavell's theory of metacognitive knowledge by showing task and strategy variables are consistent predictors of performance as is cognitive aptitude. Further research maintains strategy instruction as an important means of conveying metacognitive knowledge. Students without explicit strategy instruction do not perform as well as those who do receive instruction. As students acquired more experience they are more able to analyze strategy and task variables. In addition, the research suggests that strategy instruction, strategy use, and aptitude predict how successfully an individual can be at a variety of cognitive tasks.

Metacognitive Control

This section reviews the theoretical underpinnings of metacognitive control. Research in this area has focused on students' use of feedback prior to or during a task. The feedback may be self-generated or provided by another individual. Self-generated feedback is considered on-line monitoring. Recent research identifies individual differences in monitoring ability. A second area of on-line monitoring is the examination of feeling-of-knowledge judgments. Two theories that may explain what the feeling-of-knowing judgment represents are presented.

In his original conceptualization, Flavell did not elaborate on metacognitive control to the same degree as he did on metacognitive knowledge. Other researchers have identified three metacognitive control components: 1) planning, 2) evaluation, and 3)
regulation (Jacobs and Paris, 1987). Planning is the mental coordination prior to initiating a task. Evaluation is an assessment of the individual's understanding of the information. Regulation is the on-line monitoring during a task. It includes strategy assessment and tracking progress towards the goal of the task. This taxonomy has been adopted by others (Brown, 1987; Kluwe, 1987).

Schraw and his associates suggested that metacognitive regulation consists of planning, monitoring, and evaluating (Schraw & Dennison, 1994; Schraw & Moshman, 1995; Schraw, 1998). Planning refers to the selection of the appropriate strategy and allocation of cognitive resources. Monitoring involves analysis of present performance and comparing current progress to previous experiences (Chi, 1987). Evaluation is a process in which one determines the utility of a chosen strategy after the task has been completed. Much of the research in self-regulation has focused on metacognitive monitoring. Self-regulation is the ability to scrutinize cognitive activity and make changes as needed to meet a specified goal.

Schunn, Lovett, and Reder (2001) provided feedback in an inductive reasoning task. Subjects were instructed to build a stick of a specified length from any combination or number of three different length sticks. Two strategies—overshoot and undershoot—were identified for each subject. Strategy use was evaluated over 70 trials. In ten trials either strategy would work, for 30 of the trials the overshoot strategy was more successful, and in 30 of the trials the undershoot strategy was more successful. Subjects could choose either strategy in any trial. Scores were measured as the number of moves required to reach the desired stick length. The researchers found strategy selection inversely related to working memory and inductive reasoning. Subjects with strong strategy selection used reasoning...
and working memory less. This strategy awareness contributed to task automaticity. They also found large individual differences in strategy selection. All subjects used both strategies; however, some subjects did not connect their strategy selection with their performance.

The researchers repeated the stick-building task with a new group of subjects, but without the first 10 trials where either strategy would work. The subjects were given 30 trials requiring the undershoot strategy and thirty trials were the overshoot strategy were needed. Subjects determined which strategy to employ at each trial. In this study, subjects were more perceptive as to which strategy was more effective. Any bias for one strategy over another was not found. Strategy awareness was related to strategy adaptivity. Subjects who recognized the required change in strategy effectively made the change immediately.

Finally, four strategies where given to subjects to choose from and they were given a secondary task to do simultaneously with the stick building task. Subjects were placed in either a high-load condition or a low-load condition for the secondary task. The high-load condition required subjects press the “z” key on a computer console when they heard a letter (a or b) that was the same as the last letter or press the “x” key when the letter was different from the previous letter sound. The low-load condition required subjects to press the “z” key when they heard an “a” sound and press the “x” key when they heard the “b” sound. Working memory, as measured by doing two tasks at time, did not play a large role in strategy adaptivity. Performance was not hindered by either the high-load or low-load working memory condition. The relationship between awareness and adaptivity remained strong.
This research provides evidence for the importance of strategy selection in a task. Theoretically, strategy awareness was considered to be a part of strategy adaptivity. While the researchers expected strategy implementation would be more implicit, implicit strategy use is not necessarily disconfirmed in this work.

Despite the results of this study, Schunn, Lovett, and Reder (2001) argue that regulatory processes are automatic and implicit. The argument is that during learning the information to be encoded is not consciously evaluated. To support this contention, Reder (1996) presented subjects with a list of new names over several presentations. Subjects had to identify which names were famous and names which were new names. The subjects demonstrated high false alarm rates when the new names were embedded among famous names. Similar results were obtained in the failure to detect spelling errors after multiple exposures to misspelled words. Reder argues that the information was encoded as is and not evaluated for errors. In other words, subjects were engaged in implicit encoding.

The same approach has been applied to strategy selection. Subjects were given two choices in strategy selection in solving two-digit by two-digit arithmetic problems (Reder & Ritter, 1992). Subjects were given a large number of problems to solve as training. They were then tested and made feeling-of-knowing judgments on their answers. The subjects could attempt to retrieve the answer from memory or compute the answer. These were the two strategy choices. Strategy selection was based on familiarity with the features of each problem. The more familiar the features, the more likely the retrieval strategy was attempted. Computation was used when the subject felt less familiar with the problem. Similar results were obtained with reading comprehension strategies (Reder, 1987). Subjects were given passages to read, then tested. If test items were similar to the passage 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
a direct retrieval strategy was employed. However, if the questions were less like the passage then a plausibility strategy was used. The plausibility strategy was the relating of information in the passage to the inferences made by the question. These results were interpreted as strategy selection being automatic and implicit. This was given as evidence for the lack of conscious awareness of regulatory processes.

**Monitoring Ability**

Monitoring refers to the ongoing assessment of achievement towards a goal and the effectiveness of the strategy being used (Devolder, Brigham, & Pressley, 1990). Monitoring includes personal judgments of familiarity with a task and cues regarding the correct answer (Koriat, 2000; Schwartz & Metcalfe, 1992). Metacognitive monitoring compares recent outcomes to create performance expectations. Such judgments are used to allocate cognitive resources for the current task, such as length and depth of memory search. Pressley and Ghatala (1990) considered cognitive monitoring to be the foundation of self-regulated thinking. Initially this type of assessment is thought to occur consciously. With time it becomes automated and conscious involvement is not required (Delclos & Harrington, 1990; Narens, Jameson & Lee, 1994; Pressley, Borkowski, & Schneider, 1987).

Recent research has considered individual differences in on-line monitoring ability. Swanson (1990), for example, divided children into four groups: 1) high intelligence, high metacognitive ability; 2) high intelligence, low metacognitive ability; 3) low intelligence, high metacognitive ability; and 4) low intelligence, low metacognitive ability. Each group was given two Piagetian tasks to solve. The first was an analytical task in which a colored...
liquid was to be derived from the combination of two out of four colorless liquids. The second was a pendulum task, in which the child had to determine how to increase and decrease the speed of a pendulum by varying the weight and length of a string. Students who rated themselves as having high metacognitive skills outperformed others, despite ability. Metacognitive skills appeared to compensate for aptitude.

Corkill and Koshida (1993) found that college students with high metacognitive strategy awareness, as measured by the Metacognitive Awareness Questionnaire, had more accurate estimates of their performance with regard to aptitude tests of verbal reasoning and computation. Knowledge about cognition was related to accurate monitoring of test performance. Slife and Weaver (1992) also found a positive relationship between metacognitive monitoring and mathematical performance in depressed and non-depressed subjects. Depressed college students showed lower monitoring than the non-depressed students. Cognitive ability was not correlated with metacognitive monitoring.

Schraw (1994) tested monitoring ability in college students in reading comprehension. He found that subjects with high cognitive monitoring skills performed better and were more confident than subjects with low monitoring ability. Schraw suggested that high monitors produced better on-line, self-correcting feedback during the test. Contrary to other studies (Corkill & Koshida, 1993; Koriat, Lichenstein & Fischof, 1993; Metcalfe, 1998; Pressley & Ghatala, 1990;), poor monitors appeared to realize that their metacognitive abilities were poor. In other words, their questionnaire scores matched their confidence judgments.

Schraw and colleagues (Schraw, 1994; Schraw, Dunkle, Bendixen, Roedel, 1995; see also Schraw, 1997; Schraw, 1998) have examined the relationship between global
monitoring and domain specific monitoring. Global monitoring is a general ability applicable across all situations. Domain specific monitoring is the ability to monitor strategies specific to that task or domain. The domains examined were reading comprehension, spatial ability, and math ability. Monitoring proficiency in one domain was related to the monitoring ability in other domains. Proficiency remained consistent after domain-specific ability was controlled. When domains were matched by difficulty in a principle component analysis, a general monitoring skill was found (Schraw & Nietfeld, 1998). It was hypothesized that children first learn domain-specific regulatory skills and then generalize and adapt these skills to new situations. Schraw (1997; Schraw & Graham, 1997) proposed that experience was what influenced the movement from domain-specific regulation to more global monitoring.

**Feeling-of-Knowing**

One of the more reliable measures of cognitive monitoring has been the feeling-of-knowing judgment (FOK) (Nelson, 1984; Leonesio & Nelson, 1990). First identified by Hart (1965), a feeling-of-knowing judgment occurs when an individual can recognize information that could not be recalled initially. For example, an individual would be given material to learn, such as noun pairs or trivia. When an item was not recalled at testing, the individual was asked how confident they were that they would recognize the correct answer. Several choices were presented from which the subject could choose the appropriate answer. Feeling-of-knowing judgments are indicative of performance on the recognition task. When judgment is high, the correct answer is typically recognized.

A number of elements have been identified which affect the accuracy of feeling-of-knowing judgments. In one study, subjects were divided into four groups (Nelson,
Leonesio, Landwehr, & Narens, 1986). Each group was presented with questions that no one could answer in a free recall situation. Questions were presented in random order. The first group was shown the answer to each question for one second. The second group was given a perceptual identification task, in which the correct answer to each question was flashed on a screen at increasing durations. The subject repeated the answer when they could identify it from the flashes. The third group was given a four-choice recognition test of the same 12 questions. The last group was given an eight-choice recognition test. FOK judgments were compared to the normative item difficulty and normative feeling-of-knowing for all participants. In all groups, individual FOK judgments were more accurate at predicting performance than the normative FOK judgments, but were not as accurate as item difficulty. Greater accuracy was found in the relearning group—the first group that was shown the answers for one second—and the two groups that received the forced-choice recognition tests. These results suggested that the accuracy of feeling-of-knowledge judgments are dependent on how cues are presented.

Others have demonstrated that FOK judgments increase with self-generated feedback (Glenberg, Sanocki, Epstein & Morris, 1987). Glenberg, et. al. (1987) conducted a study in which comparisons were made on confidence judgments between domains subjects were familiar with and domains containing unfamiliar knowledge. Participants were more confident with familiar information than with unfamiliar information. In a familiar domain, comparisons were made between inference tests and verbatim tests. While subjects in the inference group had higher familiarity than those who took the verbatim test, confidence judgments during the testing did not differ. Subjects were also tested when feedback was given in the form of a pre-test. After the pre-test, subjects made
confidence judgments with respect to how they would perform on the upcoming test. Subjects who took the pre-test had more accurate confidence judgments than those who did not take a pre-test.

Koriat and colleagues tested subjects on general knowledge asking for confidence judgments along with their test answers (Koriat, Lichenstein & Fischof. 1980). In the first condition, subjects only chose one answer from four choices. In the second condition, subjects had to write down reasons as to why the choice selected was the correct alternative, as well as any reasons why the choice was incorrect. More accurate judgments occurred in the justification condition. The results suggested that confidence judgments could be more accurate when individuals compare contrary and favorable information. The authors suggested contrary evidence is often not evaluated when confidence judgments are made.

In a later study, rewards were given when correct answers were selected in a general knowledge test (Koriat & Goldsmith. 1998). Participants were tested in both free recall and recognition tasks. Subjects performed better when rewards were given. In fact, subjects in the free recall task who were rewarded performed better than those who were rewarded in the recognition task.

Feeling-of-knowing judgments did not, however, reflect performance in problem-solving tasks (Metcalfe. 1986). First, subjects were given a free recall general knowledge test and a feeling-of-knowledge judgment was made. Subjects then completed an eight-option recognition test. In a second task, subjects were given insight problems to solve. Subjects made confidence judgments and were then retested on the same problems. When the two tasks were compared, the general knowledge FOK judgments were more accurate.
than the problem-solving judgments. The FOK judgments associated with the problem-solving task were very poor. An additional analysis indicated that subjects with high confidence judgments tended to have better performance scores.

Two theories have been developed in an attempt to explain what feeling-of-knowing judgments represent. Target availability theory states that the judgment represents an incomplete retrieval of information in storage (Nelson, 1984; Nelson & Narens, 1990; Nelson, 1996). It is an established theory with many studies presented in its defense. The newer theory, cue availability, suggests that a feeling-of-knowing judgment does not represent retrieval from storage; rather it is a judgment of familiarity of the information for the individual. Many recent studies provide support for the cue familiarity theory. A more in-depth examination of each possibility follows.

Target Availability

The first theory of metacognitive control explaining feeling-of-knowing judgments was target availability (Nelson & Narens, 1990; Narens, Jameson, & Lee, 1994; Nelson, 1984; Nelson, 1996; Plude, Nelson, & Scholnick, 1998). Since a feeling-of-knowing judgment occurs when an individual knows the answer but cannot produce it; retrieval from memory must be based on how accessible that information is. When the target is accessible, it is available for recall and the answer is produced. When it cannot be recalled, recognizing the correct answer among a list of choices provides evidence that the information had been learned and a highly confident judgment was correct.

recognized as being in storage, but it cannot be retrieved without cues. When a confidence judgment is made, it is a determination by the individual of the contents of their memory. The higher the ranking of the judgment, the more confident the individual is that they know the information without retrieving that information.

One of the first studies examined feeling-of-knowing-judgments on information in long-term memory (Nelson, Gerler, Narens, 1984). Subjects were given questions regarding general knowledge in a free recall task. FOK judgments were made on questions not answered correctly. Then a four-choice recognition task was given. After the initial recall task, answers to the questions were presented briefly via a tachistoscope to one group of subjects. A second group did not view the answers. Subjects had more accurate FOK judgments when answers were shown through the tachistoscope compared to those who had the recognition test only. In the next phase, one group of subjects was given a five second glimpse of each correct answer. Another group had the answer shown in a perceptual identification task, where the answer was flashed on a screen by increasing durations until the subject identified that answer. After a distracter activity all subjects were given another recall task. Subjects had greater feeling-of-knowing judgments in the relearning condition—those who were shown the answers for five seconds. Greater response latencies were identified in the incorrect responses. This suggested that the amount of time needed for recall was determined by what an individual felt they knew. The authors theorized that there is an internal monitor that is used which compares the question to what is in memory.

In a later study, when the answer to a general knowledge question could not be recalled, the subjects were given one of four tests regarding their feeling-of-knowing.
Feeling-of-knowledge judgments were compared between word pairs that were related to word pairs that were unrelated (Carroll, Nelson & Kirwan, 1997). Subjects were given 40 word pairs to learn. Twenty of the word pairs were shown twice; these word pairs were weakly related: for example, student-paper. The twenty unrelated word pairs (such as, engine-disease) were shown eight times. One group was tested on all word pairs after a two-week delay and the other group was tested after six weeks. FOK judgments were highest for word pairs in the unrelated pairings condition. The unrelated group had more accurate judgments. There were no differences between confidence judgments and the amount of delay in testing. However, the two-week delay group did have higher performance scores than the six-week delay group. The authors suggested FOK judgments are influenced by the degree to which something had been learned. Apparently participants did not take into consideration the effect of length of delay between the learning phase and the testing phase when making confidence judgments.

This theoretical framework explains feeling-of-knowing judgments in terms of an interaction between cognitive and metacognitive activity (Nelson & Narens, 1990; Nelson, 1992; Nelson, 1996; Plude, Nelson, & Scholnick, 1998). Cognitive activity is considered
to occur at the object level. The object level is the basic level where all reasoning, attending, encoding, and retrieval processes occur.

As opposed to the object level, the meta-level is responsible for metacognition, including planning and evaluating. The meta-level contains a model in which comparisons can be made with the current task. Information moves in a circular fashion between the object level and the meta-level and then back to the object level. Monitoring is the flow of information from the object level to the meta-level. Monitoring consists of comparing what is occurring at the object level against the model within the meta-level. Feeling-of-knowing judgments are a part of monitoring. In contrast, control is the flow of information from the meta-level to the object level. Control includes adjustments in strategy or resource assignment at the object level when the information does not match the meta-level model.

Target availability identifies three stages of metacognitive monitoring during learning. In the first, or acquisition, stage the ease of the task and what portion has been learned is monitored. The retention stage (stage two) focuses on the type of processing and amount of study time available. The last stage, retrieval, considers which strategy should be used to retrieve the information. It is at stage three, or at retrieval, that measures of confidence in knowing the information may be made. Feeling-of-knowing judgments can occur during acquisition and retrieval.

Recently, target availability theory has been challenged. Opponents claim that feeling-of-knowing judgments indicate familiarity with the information, not the ability to retrieve information. Instead, these researchers believe that the cue-familiarity theory provides a better account of feeling-of-knowing judgments.
Cue Familiarity

Cue-familiarity states that a feeling-of-knowing judgment is a heuristic identifying whether information is novel or not and the judgment is made without explicitly attempting to access information in long-term memory (e.g. Kamas & Reder, 1995; Metcalfe, 1993; Miner & Reder, 1994; Reder & Ritter, 1992; Schwartz & Metcalfe, 1992). FOK judgments are thought to represent a confidence ranking of how familiar the individual is with information related to the question.

As indicated previously, Glenberg and colleagues (Glenberg, Sanocki, Epstein, & Morris, 1987) found poor FOK judgments on reading comprehension tests when judgments focused on domain familiarity. This was interpreted to mean that individuals used familiarity with the cues in the text rather than specific text information. The authors demonstrated that judgments could be improved with feedback in the form of a pretest that is closely related to the final test.

Koriat (1993) had subjects memorize four-letter nonsense strings and recall as many letters as possible. Subjects then made FOK judgments and took a recognition test on the letter strings. Koriat found higher feeling-of-knowing judgments when more letters of the string were reported. Next, the same task was performed with various lengths of time between the recall and recognition tasks, in which subjects were engaged in a distracter activity. Retrieval was more accurate with the shorter interval. In addition, more information was recalled with the shorter interval. FOK judgments were related to the accessibility of the information. Accessibility is based on the amount of information encoded and the duration between encoding and retrieval. FOK accuracy is dependent on the accessibility of that information. Koriat (1993, 1994; Metcalfe, 1998) suggested that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
subjects may have a high level of confidence when answering a question and yet their response may be completely inaccurate.

In a later study, Koriat (1995) tested subjects using a general knowledge test. The subjects had to complete a recall test, then a recognition version. When subjects had high feeling-of-knowing judgments, they tended to be more accurate in the recognition task. When examining errors, high feeling-of-knowing judgments were associated with errors of commission more than errors of omission. In relation to the recall task, higher FOK judgments were associated with the wrong answer. These results were thought to suggest that feeling-of-knowing judgments do not reflect target retrieval, only target accessibility.

Metcalfe, Schwartz, and Jouquim (1993) suggested that feeling-of-knowing judgments do not reflect memorability; rather, feeling-of-knowing judgments were indicative of the number of presentations of a cue. For example, in one study subjects had to recall four lists of twelve word pairs in a paired-associate task. They were placed in one of four conditions: 1) learn the same list twice (A-B; A-B), 2) the second list had the same words in different pairings from the first list (A-B'; A-B), 3) the second list had new words paired with words from the first list (A-D; A-B), and 4) neither list had any similar words or word pairings (C-D; A-B). The second and third lists were more likely to result in interference. In other words, the mixture of old and new words in word pairings would make retention more difficult. (For a thorough review of interference literature see Dempster, 1995 or Dempster and Corkill, 1999). If the participant could not recall the word paired with its cue, an FOK judgment was made followed by an eight-choice recognition test. Subjects in the A-B; A-B condition had the highest recall, followed by A-B'; AB, A-D; A-B, and C-D; A-B. Subjects in this last condition were statistically less
confident than subjects in the other three groups. FOK judgments reflected the number of presentations of the word pairs. The results were the same in the recognition test phase, regardless of the difficulty due to interference from the distracters or cued recall instead of a recognition test.

Reder and Ritter (1992) used feeling-of-knowing judgments as an initial confidence judgment. When individuals were presented with a computational problem, they had to rate how confident they were that they would recognize the correct answer before solving the problem. The results suggested that feeling-of-knowing judgments were best interpreted as an initial judgment of recognition, not a trace into memory storage. In other words, feeling-of-knowing judgments served as a gauge of how familiar the individual was with the question.

Reder (1987; 1996) theorized that choosing a specific strategy was based on the demands of the task. In one such study (Reder, 1987), subjects were given stories to read. They then had to rate statements regarding the plausibility of that story. Response latencies to each plausibility statement were measured. When the questions were more similar to the sentences in the story, a direct retrieval strategy was used. When more time or reasoning was required, then a plausibility strategy was used. Subjects did change strategy when the different types of questions were interspersed within the same story. The strategy selection appeared to be automatic and without conscious control. Reder suggested that familiarity leads to direct retrieval from memory and shorter response latencies. In a related finding, Reder and Ritter (1992) determined that the use of heuristics, such as feeling-of-knowing judgments, could be used to provide a quick gauge of subject familiarity with a task.
Several researchers have concluded that individuals tend to be overconfident in their FOK judgments (e.g., Glenberg, Sanocki, Epstein, & Morris, 1987; Koriat, Lichenstein & Fischof, 1993; Metcalfe, 1998; Pressley & Ghatala, 1990). When asked for confidence ratings in the feeling-of-knowing paradigm, subject ratings were consistently higher than the ability to recall the learned data. Metcalfe (1993, 1998) suggests such overconfidence is the result of a heuristic predicting future performance based on past experience, tailored to the current question. In other words, subjects believe their answer is good enough and correct (Metcalfe, 1986b). Over-optimism also contributes to overconfidence. Being so optimistic, however, may lead to missing relevant aspects of the task, such as looking for contradictory information in a passage (Epstein, Glenberg, & Bradley, 1984). In addition, Corkill and Koshida (1993) found those who assess themselves as high metacognitive monitors were more accurate at assessing their abilities, but still overconfident.

In summary, considerable research has been conducted in the field of metacognition on feeling-of-knowing judgments and monitoring. Feeling of knowing judgments are one measure of monitoring and metacognitive control. There remains an active debate between theoreticians of the make-up of FOK judgments. In a parallel area of study, monitoring ability individual differences have been identified. Monitoring has been associated with the acquisition of domain specific information. The present study proposed to tie monitoring research more closely with FOK judgments. There should be similar individual differences in FOK judgments that have been identified in the metacognitive monitoring research.
The Present Study

Despite all of the research on metacognitive monitoring and feeling-of-knowing judgments, little attention has been given to whether feeling-of-knowing judgments are a part of the monitoring continuum between those with high metacognitive ability and low metacognitive ability. It should be expected that there are individual differences in FOK judgments as a reflection of monitoring ability. In the present study these individual differences were measured by two different tasks: a paired-associate task to represent judgments related to new learning and a test of general knowledge representing judgments from prior learning. To assess monitoring ability, the Metacognitive Awareness Inventory was used.

Metacognitive Awareness Inventory

The Metacognitive Awareness Inventory (MAI) was developed to measure both metacognitive knowledge and control (Schraw & Dennison, 1994). It is a 52-item questionnaire. Internal consistency scores range from 0.88 and 0.93 (Schraw & Dennison, 1994). Each question was written to represent one of six sub-components of metacognition, (declarative knowledge, procedural knowledge, conditional knowledge, planning, monitoring, and evaluation). However, factor analysis did not recognize these six components. Two factors were identified in confirmatory factor analysis according to the theoretical basis of metacognitive knowledge and metacognitive control. Because the factors were identified through confirmatory factory analysis instead of a data driven investigation, this inventory may lack the power of exploratory factor analysis. Indeed, six of the test items had cross-loadings and two items had no loading whatsoever.
To determine the validity of the MAI, the inventory items were compared to confidence ratings and performance on reading comprehension tests (Schraw & Dennison, 1994). Subjects were placed in one of three groups according to confidence ratings. The high and middle confidence groups had significantly higher scores on Factor 1—metacognitive knowledge—than the low confidence group. There was no difference between groups on Factor 2—metacognitive control—or total MAI score. When subjects were divided into three groups by performance and compared to the MAI, both factors—metacognitive knowledge and metacognitive control—were significantly related. High performance scores corresponded to high scores on both metacognitive knowledge and control on the MAI. Subjects with low performance scores also had low scores in metacognitive awareness and control. Factor 2, or metacognitive control, was expected to be predictive of FOK judgments (Schraw, personal communication February, 2002).

Others have used the Metacognitive Awareness Inventory in research related to the use of metacognition in domain specific instruction. For example, resident physicians in anesthesiology were compared on several inventories in metacognitive ability, domain knowledge, gender, and motivation (Plants, 2000). There were significant correlations between metacognition, domain knowledge, and motivation. Furthermore, metacognition was influential with respect to medical intern’s performance after prior knowledge was partialled out in statistical analysis. Interns with better metacognitive skills, as identified by the MAI, were more likely to be successful.

Other studies have investigated using hypertext links in teaching metacognitive skills. Stimson (1999) found the Metacognitive Awareness Inventory predicted learning using hypertext. Hypertext is a highlighted phrase in an electronic document that leads to
more in-depth or related material. In this case, when the hypertext was accessed, questions regarding metacognitive monitoring were presented. Subjects were placed in a group that either read hypertext material or a group that read a linear text. Subjects with the hypertext passages had better performance scores. In addition, the MAI had high predictive value for subjects in the hypertext, but not the linear text group. Presumably individuals with high MAI scores were better at self-regulation during learning. These subjects were more likely to realize when they would need to access the hyperlinked text, and were more inclined to use such hyperlinked text. The additional information gained from the hyperlink text would then result in improved performance on a recall test. Individuals with less awareness, as indicated by the MAI, would be less likely to utilize the hypertext links, because they would not realize they did not understand the material. Subjects who read linear text appeared to function as a control group for this study. In other words, MAI scores are not related to performance scores because even when a subject realizes they need additional information in order to understand the text, no additional information was available.

When metacognitive cues were embedded in web pages that taught metacognitive skills, cognitive regulation cues were successfully conveyed whereas cognitive knowledge cues were not (Moore, 2000). In Moore’s study, subjects completed the MAI and a computer-based instructional module. Each participant was placed in one of three groups: 1) knowledge of cognition cues, 2) regulation of cognition cues, and 3) no cues. In the knowledge of cognition group, cues related to metacognitive knowledge were embedded in the web page. In the regulation of cognition group, cues related to metacognitive control were embedded in the instructional web pages. The regulation of cognition group had
better performance scores than the other groups. There were no significant differences between groups on the MAI.

This study demonstrated the importance of metacognitive control cues in instruction. Moore did not find any difference between metacognitive knowledge and metacognitive control groups. This should not be surprising since the MAI is composed of items that reflect both metacognitive knowledge and control.

Metacognitive awareness had little effect with veterinarian students who were taught using case studies (Law, 1999). In this study, veterinarian students were given several case studies in veterinarian pharmocokinetics. Students were placed into one of two conditions. subjects either had use of a provided pharmocokinetic model or they had to assemble a model of their own. Metacognitive awareness was measured with the MAI. The group that assembled their own models performed better on the case study test than subjects with a provided model, but the MAI had no predictive value for either group.

**Need for Cognition**

The Need for Cognition scale was developed by Cacioppo and Petty (1982). Need for cognition is a stable trait reflecting the willingness or motivation of an individual to become involved in effortful cognitive activity (Tidwell, Sadowski, & Pate, 2000). Individuals high in need for cognition tend to engage in and enjoy thinking (Cacioppo & Petty, 1982). They also tend to organize, elaborate, and evaluate information they have been exposed to. Individuals who measure high on this scale tend to enjoy complex tasks. Individuals high in need for cognition also report less frustration and mental discomfort during cognitive tasks. Need for cognition has been shown to be unrelated to gender, sex
role attitudes, and sociability (Cacioppo & Petty, 1982; Osberg, 1987). A factor analysis resulted in one factor, need for cognition. Split-half reliability was reported at 0.87.

The Need for Cognition scale is positively correlated with verbal intelligence (Cacioppo, Petty, Kao, & Rodriguez, 1986), ACT scores (Cacioppo and Petty, 1982), domain knowledge (Tidwell, Sadowski, & Pate, 2000), and verbal ability (Kardash & Noel, 2000). Positive relationships have been found between the Need for Cognition scale and measures of self-esteem, personal attitudes, and attention (Osberg, 1987). Negative associations have been found between the Need for Cognition scale and self-consciousness and social anxiety.

Two other forms of the Need for Cognition scale have been developed. The short form developed by Cacioppo, Petty, and Kao (1984) retained 18 of the original 45 items. The reliability between the original and short form was 0.95. Cronbach alpha on the short form ranged between 0.86 and 0.95 (Cacioppo, Petty & Kao, 1984; Sadowski, 1992; Sadowski & Gulgoz, 1992). Test-retest reliability over a seven-week span was 0.88. The short form also has been positively correlated with ACT scores and grade point average (Tolentino, Curry & Leak, 1990). Preference for reading material, magazines, fiction and nonfiction books, have been positively correlated with need for cognition using the short form.

Ahlering and Parker (1989) have suggested that individuals with a high need for cognition process information differently than those with a low need for cognition. In a study, subjects were compared on their Need for Cognition scale scores and response latencies in ranking individuals according to a list of adjectives read to the subject. Response latencies were not significant between subjects with high NFC scores and those
with low NFC scores. When looking for a primacy effect, subjects with high NFC scores did have lower mean primacy scores than those with low NFC scores. In other words, subjects with high need for cognition were less susceptible to primacy effects. Incredibly the authors interpreted this to mean that subjects with high need for cognition process information differently than those with low need for cognition.

Miner (1992) found that individuals with a high need for cognition outperformed those with a low need for cognition in conceptually driven encoding and retrieval tasks. Subjects with low need for cognition did better at data-driven encoding and retrieval tasks. Kardash and Noel (2000) gave subjects an essay, followed by recall and recognition tasks. Subjects with higher need for cognition scores recalled more than those with low need for cognition scores. These studies suggest that those with low need for cognition may make use of information and cognitive resources differently than those with a high need for cognition. Unfortunately, the methods and variables used in these studies are insufficient for identifying the extent to which information may be processed differently or to identify individual differences in cognitive resources used to complete these tasks. Swanson (1990; Swanson & Trahan, 1996) has shown that individuals with high metacognitive ability have better performance scores than those with low metacognitive ability. The similarities between metacognitive ability and need for cognition may indicate a relationship between the two constructs.

Tanaka, Panter, and Winborne (1988) developed a modified Need for Cognition scale using 34 of the original items. Through exploratory factor analysis three subscales were identified: cognitive persistence, cognitive confidence, and cognitive complexity. Cognitive persistence reflects how focused one believes oneself to be in a cognitive task.
Cognitive confidence represents how successful a person expects to be in correctly completing a task. Cognitive complexity reflects the degree to which an individual enjoys the intricacies of a task. Coefficient alpha scores for internal consistency on the total modified form range between 0.77 and 0.89 (Waters & Zakrajsek, 1990). Subscale reliability scores, also measured by coefficient alpha, were: cognitive persistence between 0.68 and 0.81; cognitive confidence between 0.59 and 0.68; and cognitive complexity between 0.57 and 0.71. Small but significant positive correlations were found between the modified Need for Cognition form and ACT scores as well as grade point average.

It has been demonstrated that metacognition and intelligence are not highly related (Swanson, 1990; Swanson, Christie, & Rubadeau, 1993; Swanson & Trahan, 1996) while need for cognition and intelligence is (Cacioppo and Petty, 1982; Cacioppo, Petty, Kao, & Rodriguez, 1986; Kardash & Noel, 2000; Tidwell, Sadowiski, & Pate, 2000). Still, certain subscales in the Need for Cognition scale, specifically cognitive persistence and cognitive confidence, may be related to metacognitive monitoring processes. Metacognitive knowledge of person and strategy variables may have an effect on how confident and persistent an individual is at a cognitive task. If a person has high awareness of their cognitive abilities they should be highly confident and persistent in cognitive tasks. This has yet to be investigated.

One problem with measuring metacognitive monitoring is that the questions typically asked on monitoring scales may reinforce strategies or give subjects new ideas on how to proceed in a cognitive task. One solution to this problem would be to provide enough questions that the purpose of the study is sufficiently camouflaged. The Need for Cognition scale was used for two reasons: 1) to investigate the relationship between Need
for Cognition and metacognitive awareness, and 2) to mask the nature of the items of the Metacognitive Awareness Inventory.

The goal of this study is to assess the individual differences between subjects with high metacognitive monitoring ability and those with low metacognitive monitoring ability in retrieval from memory and in encoding data using feeling-of-knowing judgments. The hypotheses in this study were:

1. Individuals with high metacognitive control scores from the Metacognitive Awareness Inventory would be more accurate in confidence ratings while those with low metacognitive control scores would have greater inaccuracies in confidence ratings in a recall/recognition task.

2. Individuals with high metacognitive monitoring scores would have significantly greater accuracy in confidence at recall in a paired-associates task than those with low monitoring scores.

3. There would be positive relationships between the Need of Cognition scale and the Metacognitive Awareness Inventory. The monitoring scale in the Metacognitive Awareness Inventory would be positively related to the NFC subscales cognitive persistence and cognitive confidence.
CHAPTER 2

METHOD

Subjects

Participants consisted of 75 undergraduate students from the University of Nevada, Las Vegas recruited from the Educational Psychology Department Subject Pool during the Spring 2002 semester. There were 11 males and 64 females. The average age was 25.4 years. Each volunteer received research credit in partial fulfillment of a course requirement. It was assumed this sample would be relatively unbiased and random. Informed written consent was obtained prior to participation from each individual, according to University of Nevada, Las Vegas protocols.

All participants were tested together in a typical classroom in groups of 10 to 15. Participants were divided into three groups according to ability as measured by the second factor, metacognitive control, of the Metacognitive Awareness Inventory. The 25 participants with the lowest Metacognitive Awareness Inventory scores were assigned to the low monitoring group. The 24 individuals with the highest scores on the Metacognitive Awareness Inventory were assigned to the high monitoring group. The remaining 26 constituted the middle group.

Materials

Metacognitive Awareness Inventory.
The second inventory was the Metacognitive Awareness Inventory. This consisted of a 52-item self-report inventory developed by Schraw and Dennison (1994). The Metacognitive Awareness Inventory had six categories separated into two factors, metacognitive knowledge and metacognitive regulation. Coefficient alpha scores for internal consistency range between 0.88 to 0.93. The second factor, metacognitive regulation was used to discriminate individuals according to monitoring ability.

Need for Cognition scale.

This study used the modified Need for Cognition scale developed by Tanaka, Panter, and Winterborne (1988). The modified form has only 34 of the original 45 items. The main purpose of the Need for Cognition Scale was to mask the metacognitive monitoring measurement. It was hoped this would prevent participants from acquiring and using information from the Metacognitive Awareness Inventory. The modified form has three subscales identified through factor analysis: cognitive persistence, cognitive confidence, and cognitive complexity. The Cronbach alpha for the modified Need for Cognition scale ranged between 0.77 and 0.89. The alphas for the subscales ranged for cognitive persistence between 0.68 and 0.81, for cognitive confidence between 0.59 and 0.68, and for cognitive complexity between 0.57 and 0.71.

Recall task.

The recall task consisted of 27 items selected from Nelson and Narens (1980) 300 general knowledge questions. There were 25 test items plus 2 practice items. These questions covered subjects such as American and World History, Science, Sports, Literature, and Entertainment (sample question: What is the highest mountain
in South America? answer: Aconcagua). Twenty-five items were selected at random where the probability of recall was 0.496 or less, according to the original normative data. The recall probability mean of all twenty-five items was 0.139. This criterion was chosen to allow moderate recognition answers without having high recall. Five questions were excluded from consideration because males had a significantly higher probability of recall than females. A pilot study maintained a gender difference on these items. Four other questions were excluded because of changes in the last twenty years from when the original data was tested (e.g., What is the capital of Czechoslovakia? This item was excluded because Czechoslovakia no longer exists.)

The final pool from which the 25 questions were selected was 191 items.

Answers were written on a separate sheet. The answer sheet consisted of a number corresponding to each test item, a line for the answer, and confidence ratings in intervals of twenty—0, 20, 40, 60, 80 and 100. The confidence intervals were based on Reder and Ritter (1992; see also Koriat, 1995; Nelson, 1996; Carroll, Nelson & Kirwan, 1997). The questions and answer sheets appear in Appendix I.

Recognition task.

The recognition task used the same items and procedure in the recall task; however, the items were presented with five answer choices: four distracters and the correct response. The test items were randomly re-ordered. The answer sheet reflected the difference in the task. There were five letters, A through E, representing each answer choice. The participant circled the corresponding letter. Corresponding confidence intervals of twenty appeared before each test item. The multiple-choice questions with the distracters and the answer sheet appear in Appendix I.

Paired-Associate Task.
The paired-associate task consisted of 20 noun-noun pairs presented orally on a previously recorded audiocassette to insure consistency in presentation. At encoding, subjects heard the noun-noun pairs at 3-second intervals. At recall, one noun was read and the participant wrote the corresponding noun. Participants had 5 seconds to recall the corresponding noun on the recall task. At the first trial, each word pair was presented. Subjects were then tested. In the second trial, the same word pairs were presented but in a different order. This ordering was established by random selection of the pairings. In the third trial, pairings included seven of the original pairs, seven of the words paired with a new word, and six novel word pairings. Each trial was recorded on a separate answer sheet. These sheets included confidence ratings prior to each response. Each list and the answer sheet may be found in Appendix I.

Procedure
The author collected all data, after approval of the thesis committee and the University of Nevada, Las Vegas Institutional Review Board for the ethical treatment of human subjects.

All participants first completed the Need for Cognition scale and the Metacognitive Awareness Inventory. This took approximately 20 minutes.

The Metacognitive Awareness Inventory was used to divide subjects into groups according to monitoring ability. The 24 students with the highest Metacognitive Awareness Inventory scores were placed in the high monitor group. The 25 students with lowest Metacognitive Awareness Inventory scores were placed in the low monitor group. The middle monitoring group was comprised of the remaining 26 subjects.
The second half of the session included the recall and recognition tasks and the paired-associate task. The recall task consisted of 25 trivia question taken from the Nelson and Narens (1980) normative data. The instructions were read aloud by the experimenter and contained the following:

The purpose of this experiment is to investigate your knowledge about general information. After you read each question you should circle how confident you are that you know the correct answer. Then, PRINT your best guess in the space provided on the answer sheet. If you are positive you do not know the answer then circle the confidence level and write NEXT on the line provided. The answer will always consist of exactly one word and will never be longer than one word. The questions differ in how easy or hard they may be to answer.

Now let's try a couple of practice questions. Move your answer sheet to one side. The first question is "What is the name of the horse-like animal with black and white stripes?" Circle how confident you are that that is the answer and write your answer in the space provided. Now look at the second question. It asks, "What is the last name of the author who wrote 'The Old Man and the Sea'?" Again circle how confident you are that that is the answer and write your answer in the space provided. Now, if you have any questions about the procedure raise your hand. The experimenter cannot answer any questions once you have begun. When you are ready turn to the next page and begin.
Individuals then completed the recall task by reading the question, circling their confidence rating, and then writing down their answers. Each individual was given unlimited time to complete this task.

The next portion consisted of the recognition task. The same twenty-five questions were presented in a new randomized order. However, five answer choices were given with the question. The same instructions were given with the modification being that the participant would circle the letter on the answer sheet that corresponds with the correct answer. Again the corresponding confidence level was requested prior to answering the question. The first two questions were examples and used as demonstration. They were not included in the analysis. Subjects were given unlimited time to finish.

The final component of the testing session was the paired-associates task. Subjects heard twenty noun-noun pairs from a pre-recorded audiocassette. At recall, one noun, the cue, was presented and each participant wrote the corresponding noun on the answer sheet. Individuals circled how confident they were at retrieving the correct noun. This process was repeated until all twenty cue-nouns had been heard and recorded with each corresponding confidence rating. The procedure was repeated for the second trial. Then a third trial list was presented in the same manner. Each participant wrote down the confidence rating and new pairing as before. It was expected the AB-ABr task would measure differences in monitoring ability and interference.
CHAPTER 3

RESULTS

The Metacognitive Awareness Inventory was scored in the manner described by Schraw and Dennison (1994). The Cronbach alpha on the MAI was 0.92. Scores on Factor 2 (metacognitive regulation) of the Metacognitive Awareness Inventory were tallied and each participant was ranked from highest to lowest, according to their scores on the inventory. The alpha for Factor 2 was 0.89. Factor 2 items are found in Appendix 1.

Groups were divided into thirds by percentile rank at 33% and 67%. There were 24 people in the high monitoring group, 26 in the moderate monitoring group, and 25 in the low monitoring group. The means and standard deviations of MAI Factor 2 scores for each group are presented in Table 1. In order to establish that the three monitoring level groups were different in terms of self-reported monitoring ability, a one-way analysis-of-variance was conducted with group (high, middle, low) as the independent variable and MAI Factor 2 as the dependent variable. The groups were significantly different, $F_{(2,72)} = 226.10$, $MSe = 40.65$, $p < 0.05$. A post hoc Tukey HSD indicated that all groups were significantly different from each other.

Traditionally the Goodman-Kruskal rank correlation gamma statistic has been used in work with feeling-of-knowing judgment ratings (Nelson, 1984). Some (Kelemen, Frost & Weaver, 2000) consider the gamma (G) the best measure of relative metacognitive accuracy. There were several issues, however, that precluded use of the Goodman-
Kruskal gamma in this study. Specifically, there were a number of confidence judgment ties between incorrect and correct responses. This resulted in a large number of gammas that were either +1 or -1, leading to ceiling and floor effects. This did not provide enough variance for analysis. Therefore, gamma was not the most appropriate approach under these circumstances (Goodman and Kruskal, 1954). Several other approaches for examining confidence exist, including Briar's index (Yates, 1990; Keren, 1991), a discrimination index (Schraw & Dennison, 1994), the bias index (Yates, 1990), and Confidence Accuracy Quotient (e.g. Keren, 1991). Two approaches were deemed appropriate for this study: 1) a bias measure (Yates, 1990; see also Schraw & Netfield, 1998) because it is considered a measure of absolute metacognitive accuracy (Kelemen, Frost & Weaver, 2000) and 2) the Confidence Accuracy Quotient (CAQ) because it reflects the degree to which a person's confidence for a correct answer exceeds their confidence for an incorrect answer (Keren, 1991; Lundeberg, Fox, & Puncochar, 1994).

The bias score is the difference between the mean level of confidence and the mean performance score divided by 100. All scores range between +1 and -1. This measures the degree to which a participant was overconfident or underconfident. Scores greater than zero represent overconfidence, while scores less than zero represent underconfidence. A score of zero represents no bias. The means and standard deviations for the bias scores for the three monitoring groups on the feeling-of-knowing task are shown in Table 2. Table 3 displays the means and standard deviations for the three groups across the paired-associates task.

A Confidence Accuracy Quotient (CAQ) reflects the degree to which the participant's confidence for a correct answer exceeds confidence for an incorrect answer.
A negative CAQ value represents higher confidence when wrong than when right. In other words, a subject is more confident about incorrect answers than correct answers. A positive CAQ value represents higher confidence when correct than when wrong; a subject is more confident about correct answers than incorrect answers. A CAQ of zero reflects an inability on the subject’s part to distinguish between instances when they are right versus when they are wrong. The CAQ means and standard deviations for the feeling-of-knowing task and the paired-associates task are shown in Tables 4 and 5, respectively.

The first hypothesis was that subjects in the high monitoring group would be more accurate in confidence ratings in the feeling-of-knowing task. This means that for the high monitoring group bias scores should be closer to zero and CAQ scores should be positive. In the feeling-of-knowing paradigm recall scores are not typically included in statistical analysis, rather they are used to eliminate items from the recognition task on a subject-by-subject basis. If a subject can correctly retrieve an answer on a recall item, that item is eliminated from the recognition task because the confidence rating for that item should be 100 percent. Thus, there were typically less than 25 items scored on confidence ratings in the recognition task for each subject. In the feeling-of-knowing task (recognition task) there was no significant difference between groups on either bias or CAQ scores as indicated by one-way ANOVA. F(2,72) = 0.80, MSe = 0.03, n.s.; F(2,72) = 0.05, MSe = 0.34, n.s. respectively.

Each monitoring group was examined to see if either bias or CAQ were significantly different from zero. With respect to bias scores, the high and middle monitoring groups were significantly distant from zero, while the low monitoring group was not (high t = 2.41, p < 0.05; middle t = 2.90, p < 0.05; low t = 1.59, n.s.).
suggests that the high and middle monitoring group were overconfident, while the low monitoring group was neither over nor under confident. The Calibration Accuracy Quotient was significant for all three groups (high $t = 3.92, p < 0.05$; middle $t = 5.00, p < 0.05$, low $t = 3.84, p < 0.05$). This suggests that they realized when they were guessing and when they knew the answer.

The second hypothesis was that the high monitoring group would have greater accuracy in confidence judgments in the paired-associates task than those who rated themselves as poor monitors. A repeated measures ANOVA was conducted for scores on the paired-associates task. No statistical differences were found between monitoring groups for either bias scores, $F(2,72) = 0.31, MSe = 0.01$, n.s. or CAQ scores, $F(2,72) = 0.77, MSe = 1.05$, n.s. Although the analysis of CAQ scores was not significant, it is worth noting that CAQ scores for subjects in the high monitoring group increased from test to test while CAQ scores for the subjects in the low monitoring group were relatively constant across the first two trials followed by a substantial decrease on the final test.

In analyzing how each group did separately, a one-way repeated measures ANOVA was conducted. The high monitoring group was significantly distant from zero across all three trials. The bias $F$ was $8.12, (1,24), p < 0.05$. The CAQ $F$ score was $241.62, (1,24), p < 0.05$. The middle monitoring group was also significant across all trials, bias $F(1,25) = 4.37, p < 0.05$, CAQ $F(1,25) = 153.47, p < 0.05$. The low monitors were not statistically significant on the bias scores ($F(1,23) = 3.88$, n.s.), but they were significant with respect to the CAQ ($F(1,23) = 102.56, p < 0.05$).

In the original factor analysis of the MAI conducted by Schraw and Dennison (1994), there were a number of cross-loadings between factors. Schraw and Dennison identified factors on theoretical rather than statistical grounds. This means that scores on
Factor 1 and Factor 2 and total MAI may be highly correlated. Due to this possibility, correlations were calculated between total MAI, Factor 1, and Factor 2. The correlation between total MAI and Factor 1 was $r = 0.78$ (p < 0.05); Total MAI and Factor 2, $r = 0.94$ (p < 0.05); and Factor 1 and Factor 2, $r = 0.53$ (p < 0.05). Because of these moderate to high correlations further analysis of bias and CAQ scores was considered appropriate.

Scores on Factor 1 (metacognitive knowledge) were tallied and each participant was ranked from highest to lowest according to how they scored on the inventory. The alpha for Factor 1 was 0.84. Factor 1 items are found in the Appendix 1. Groups were divided into thirds by percentile rank at 33% and 67%. There were 27 people in the high monitoring group, 22 in the moderate monitoring group, and 26 in the low monitoring group. The means and standard deviations of MAI Factor 1 scores for each group are presented in Table 6. In order to determine if the three monitoring groups were different in terms of self-reported monitoring ability, a one-way ANOVA was conducted with group (high, middle, and low) as the independent variable and MAI Factor 1 as the dependent variable. The groups were significantly different, $F_{(2,72)} = 99.53$, MSe = 20.57, p < 0.05. A post hoc Tukey HSD indicated that all groups were significantly different from each other.

When groups were divided using Factor 1 the results were not significant. The F ratio for bias scores on the feeling-of-knowing was $0.87_{(2,72)}$, MSe = 0.03, n.s. The CAQ F ratio was $0.82_{(2,72)}$, MSe = 0.33, n.s. For the paired-associates task the bias score the F score was $1.56_{(2,72)}$, MSe = 0.07, and F value was $0.01_{(2,72)}$, MSe = 0.01 for the CAQ score.

Groups were then divided according to the total MAI score into thirds by percentile rank at 33% and 67%. There were 25 people in the high monitoring group, 25 in the moderate monitoring group, and 25 in the low monitoring group. The means and standard deviations of total MAI scores for each group are presented in Table 7. In order to
determine if the three monitoring level groups were different in terms of self-reported monitoring ability, a one-way analysis-of-variance was conducted with group (high, middle, and low) as the independent variable and MAI total score as the dependent variable. The groups were significantly different, $F_{(2,72)} = 192.69$, $MSe = 81.49$, $p < 0.05$. A post hoc Tukey HSD indicated that all groups were significantly different from each other. The total score bias $F$ value for the feeling-of-knowing task was $0.37_{(2,72)}$, $MSe = 0.03$, n.s., and $0.41_{(2,72)}$, $MSe = 0.34$, n.s., for the CAQ score. For the paired-associates task the bias score was $F_{(2,72)}$, 0.27 $MSe = 0.01$, n.s., while the CAQ $F$ was $1.07_{(2,72)}$, $MSe = 1.01$, n.s. There were no differences in confidence judgments between ability groups.

The last hypothesis was that there would be a positive relationship between the Metacognitive Awareness Inventory and the Need for Cognition Scale. In particular, the MAI should be positively related to the Need for Cognition cognitive persistence and cognitive confidence subscales. Pearson correlation coefficients were calculated between the MAI and NFC including all subscales. The correlation between total MAI and total score on the NFC was significant, $r = 0.55$ ($p < 0.05$). The correlation between total MAI and the NFC subscale cognitive persistence was $r = 0.42$ ($p < 0.05$), and between total MAI and the NFC subscale cognitive complexity was $r = 0.51$ ($p < 0.05$). All correlations are shown in Table 8. The Cronbach alpha for the NFC was 0.50, which is much lower than previously reported (Tanaka, Panter & Winterborne, 1988). For the subscales the alpha for cognitive persistence was 0.34; for cognitive confidence alpha was 0.64; the alpha for cognitive complexity was 0.18. This hypothesis was supported by the significant correlations between the Metacognitive Awareness Inventory and the Need for Cognition subscales cognitive persistence and cognitive confidence.
CHAPTER 4

DISCUSSION

First, some possible interpretations for the first two hypotheses are presented. The relationship between the MAI and NFC is examined. The limitations of the study are then reviewed. Some limitations are inherent to all related research, for example, the reliance on self-report and social desirability. Others limitations may be due to the instruments used in this study. The last section examines what research could be done based on the results of this study and how it may fit in with previous research. This includes what improvements could be made to this study and a more in-depth look at the appropriateness of the target availability theory.

The first hypothesis was that subjects who ranked themselves high in metacognitive monitoring would demonstrate higher levels of accuracy in confidence scores on the general knowledge feeling-of-knowing task. The bias score reflects the degree of overconfidence. According to the bias scores on the recognition task, the middle monitoring group had the highest mean score, while the low monitoring group had the lowest mean score. This depicts the middle monitoring group as the most overconfident and the low monitoring group as the least overconfident. Overall, however, subjects’ confidence ratings were relatively close to zero, suggesting that subjects knew when they were selecting correct versus incorrect answers.

47

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
On the recognition task, the high monitoring group had the lowest CAQ scores, while the middle monitoring group had the highest (see Figure 2). As a whole, subjects' CAQ scores indicated that they were more confident when selecting the correct answer than when selecting an incorrect answer. In other words, they knew when they were right, but were unsure when guessing.

The selected test items were ranked between moderate to difficult in ability to answer. If the items were easy there will be little overconfidence because each item would be correctly answered. If all items were difficult there may also be little overconfidence because the subject may recognize that they do not know that answer. These questions may have been more difficult than moderate. This could account for the low bias scores. The more difficult questions may have made it easier to distinguish what was known from what was guessed. This would explain the clear distinction in the CAQ scores.

The statistical difference from zero for each group on this task suggests that on the CAQ scores each group was able to distinguish what they knew from what they did not know. While the bias scores indicate the high and moderate groups were overconfident, the low monitoring group was neither over nor under confident. This indicates the low group knew they were poor monitors. These results fit well with Schraw (1994).

The second hypothesis was that subjects who ranked themselves high in metacognitive monitoring would demonstrate higher levels of accuracy in confidence scores on the paired-associate task. When the three groups were examined by bias and CAQ scores, each group had a unique response pattern, though not significantly different, as seen in Figures 4 and 5. Subjects in both high and low monitoring group started with similar bias scores. In the second trial, the high monitors bias scores increased while the low monitors remained low. In the final trial, the low monitoring group increased their
judgments to match the high monitoring group. The middle group started with high overconfidence (the highest bias scores) and then overcorrected, having very low bias scores in the second trial. In the final trial, the middle group matched a moderate confidence score with the other two groups. Subjects were tested on the same pairings in the first two trials. In the third trial some of the same pairing were presented, some brand new pairings were presented and some of the original words were paired with different nouns. This last trial would have been more open to the influence of interference. A detailed analysis of subject performance on each type of noun-noun pair (old, new, or mixed) might yield more information with respect to this hypothesis. This, however, was beyond the scope of the present study.

With respect to the CAQ scores, all three groups performed similarly on the first trial. Higher CAQ scores reflect greater confidence in correct answers over incorrect answers. The middle group remained consistent across all trials. The high monitoring group became more confident of their correct answers across trials. The low monitoring group confidence dropped with the introduction of new material in the last trial. It appears that the low monitoring group judgments were more susceptible to the interference related to the different pairings in the third trial. A detailed analysis of the type of pairing in the third trial may shed more light on distinguishing confidence between the old pairing, the new pairings, and the mixed pairings. The paired-associate task was an encoding task and not a retrieval from long-term storage task. Because subjects were tested immediately after encoding, the paired-associate task is a working memory task. If this is the case, monitoring ability may be limited to the working memory capacity of the subject. This possibility has yet to be empirically investigated.
Each group's statistical distance from zero demonstrated the same pattern with the paired-associates task as the first task. Each group was able to distinguish what could be retrieved from what was a guess. The high and middle groups were overconfident with respect to their ability across all three trials, while the low monitoring group did not show the same overconfidence. Apparently, the low monitoring group realized that they were poor monitors.

The third hypothesis was that there would be a positive correlation between the MAI and the NFC. This positive correlation was statistically significant. The subscales on Need for Cognition, cognitive persistence, and cognitive confidence were also positively correlated with the MAI. Individuals who scored high in Need for Cognition, cognitive confidence and cognitive persistence, saw themselves as highly aware of their metacognitive activities, while those with low need for cognition saw themselves as less aware during a task. These correlations suggest persistence and confidence may be influential in metacognitive activity.

How can one account for these results? The first may be to question the validity of the MAI. Does it measure metacognitive awareness and regulation? The instrument is certainly reliable. The high correlations between the two factors and the total MAI, specifically the coefficient of 0.94 between total MAI and Factor 2, suggests only one dimension is being measured. Sperling, (aka Dennison, personal communication. May 24, 2002) suggests only one dimension is being measured, strategic knowledge. Strategic knowledge may include understanding a variety of different strategies and how to apply each strategy to different tasks. If this is the case, the results of this study indicate strategy knowledge as a related process to Need for Cognition. While monitoring may also be
involved in Need for Cognition, the evidence from this study does not support that conclusion. Unfortunately, monitoring may not have been measured by the MAI.

The result is that the groups may have been divided based on strategic knowledge than metacognitive monitoring. Those who recognized they had little strategic knowledge were not as confident in their ability. On the other hand, those with high strategic knowledge were more confident, presumably because they knew more. These particular tasks, however, may not have been conducive to the use of that strategic knowledge. This would account for the similar confidence ratings and performance scores.

The correlations between the Need for Cognition scale and the MAI indicate that strategic knowledge may be involved in whether someone enjoys cognitive activities. An individual with a high Need for Cognition may have a better understanding of strategy use in cognitive activities. They may also have more strategies available to apply to cognitive tasks. While an individual with a low Need for Cognition lacks the appreciation for strategic knowledge and its application to cognitive tasks.

Limitations

Limitations of the present study include a concern whether subjects can accurately report their own metacognitive monitoring, the effect of social desirability when responding to a self-report inventory, and the appropriateness of the MAI as a measure of metacognitive monitoring. In addition, concerns regarding high variability between groups is discussed.

The first limitation of this study is related to the use of all self-report instruments. That is, there is an underlying assumption that people know themselves well enough to provide accurate information. This may be particularly difficult when asked to assess what one is aware of while engaged in a cognitive task.
Reder (1996; 2001) has suggested that metacognitive regulation is implicit and automated, making metacognitive processes beyond the awareness of the individual. If metacognition can only be measured indirectly, then a self-report inventory will not be effective. This study may provide support for Reder's assertions. As individuals become better at implementing a strategy, the process likely becomes more automated. If strategies are beyond conscious control, then they may also be beyond monitoring. In addition, the ability to make judgments about strategies in retrieving information may be implicit. Reder (2001; Schunn, Lovett & Reder, 2001) has shown that providing feedback promotes changes in strategy, but this may also be implicit. Recognition tasks are highly familiar to college students and they may not realize that they need to evaluate their performance. Because this was a single exposure project, participants may not have been as invested had they been working toward a grade in a course. This could affect not only performance on the task itself, but also confidence. Monitoring may have been more accurate if subjects felt more invested.

Others concur with Reder's sentiments. Schwartz, Benjamin, and Bjork (1997) state that immediate metacognitive judgments are unreliable. Judgments such as in the feeling-of-knowing paradigm are based on inferences not related to the task, such as semantic related information. Confidence judgments then become inferences to whatever information can be retrieved. They have tested this hypothesis (Benjamin, Bjork, & Schwartz, 1998). In the first study, participants were quizzed on 20 easy to moderate trivia questions. They also gave confidence rankings on producing these answers at a future time. Participants were given the questions, then a test of geographical knowledge. They then wrote all the answers to the trivia test on a blank sheet of paper. Participants did not accurately predict how they would do on the free recall portion of the study. These
findings are explained in terms of using episodic memory to predict a semantic memory task.

In the second half of the experiment participants were asked to predict performance in recalling 13 words from six lists. They were tested twice: immediately after the presentation of each list, and with delay after the presentation of all six lists. Participant predictions did not reflect the primacy and recency effects found in their performance for both testing conditions. While this study may explain the results for the feeling-of-knowing task in the current study, it does not explain the paired-associate task where participants were exposed to several word pairings three times. The paired-associate task may be more reflective of working memory, not semantic or episodic memory.

The next concern is the concept of social desirability as it relates to the use of self-reports. Social desirability is an attempt to do what is considered socially correct or to do what is perceived as what the researcher wants. If college students believe they are being assessed on cognitive tasks, they may want to make themselves look better. They may take more time in assessing the task and their performance than they would in a more typical classroom setting. A college student does not want to give the impression that they are incapable of the task given to them or look deficient. This may be particularly true since the subject pool primarily consists of education majors. They may use all they have learned so as not to make an impression that they were poor learners.

The third concern is the possible presentation of critical information within the MAI. The Metacognitive Awareness Inventory identifies key points regarding monitoring, how to do it, and when to monitor. This may present new information to a research participant or provide previous metacognitive instruction in a new fashion. The subject may gained a greater understanding of metacognition during the presentation of the MAI.
and may have used that in the two tasks that followed. A more conscious attempt at metacognitive monitoring may have occurred after the presentation of the MAI providing these results.

The final concern relates to the high variability within each group on confidence judgments. One example is the calibration scores on the feeling-of-knowing task (Table 4). While the means may be sufficiently different, the standard deviations overlap. High variability in monitoring has been found elsewhere (Schraw & Dennison, 1994). However, Schraw and Dennison found differences between groups in performance as well as other indicators.

When dividing participants into three groups using the Need for Cognition scale, the same overlapping variance occurred. No statistically significant difference was found between groups in confidence ratings or performance on either task, regardless of how the groups were divided. The large mean difference on the first trial of the paired-associate task between bias scores should indicate differences in monitoring ability (see Figure 4). When looking at only the first trial, there was no statistical difference. The variance may have occurred as a by-product of this particular sample or perhaps as an artifact of subject attempts to implement information provided by the MAI.

**Future Directions**

The implications of this work may be important to educators. In particular, there is the need for an instrument that can discriminate between individuals in monitoring ability. This possibility will be reviewed. In addition, differences in theoretical outlooks are examined, such as those outlined by Flavell (1987), and the newer nomenclature used by Jacobs and Paris (1987) or Schraw (1994). Finally, cue familiarity and target availability are also reviewed in terms of this and future research.
Further investigation may show a significant interaction in the interference task between the high and low monitoring groups. In order to find differences between high and low monitors more stringent cut off points could be established. The high variance may come from a larger middle group. Perhaps individuals make inconsistently make use of metacognitive knowledge. If groups were divided into quartiles, one might find efficient monitoring in the top 25 percent and low metacognitive monitoring in the bottom 25 percent. The middle 50 percent may simply make inconsistent use of what they know.

Research in this area may provide support to those looking at differences in how information is processed (Ahlering & Parker, 1989; Kardash & Noel, 2000; Miner, 1992). Individuals with a better understanding of how they think may show differences in how they interpret a task given to them. Individuals with greater understanding of metacognition should become more comfortable with the task, while those who are less aware of their cognitive resources will be less efficient and less proficient.

Another need is an in-depth examination of instruments that purport to measure metacognition. Similar items could be pooled and then compared with related inventories to create a better understanding of metacognition and a better means of measuring metacognitive awareness. An instrument could be designed specifically to assess metacognition and discriminate between individuals. The MAI was designed as a research tool investigating monitoring specific tasks such as reading comprehension. It was not designed as an assessment tool or to discriminate between individuals.

The MAI was developed within the context of the theoretical structure developed by Brown (1987) and Paris and Jacobs (1984). The underpinnings are that metacognitive awareness consists of procedural, declarative, and conditional knowledge. Furthermore, there is great emphasis on metacognitive regulation in this theory. There are many more
items on the MAI related to metacognitive control than metacognitive knowledge. These theorists appear to focus more on regulating cognitive activity.

The original theory created by Flavell (Flavell, 1987; Flavell & Wellman, 1977) focused more on metacognitive knowledge and its underlying variables of person, task, and strategy. The Metacognition Awareness Questionnaire developed by Corkill (1993), and based on the work of Swanson, is founded in Flavell's theory. The MAQ has been helpful when assessing differences in metacognitive ability. Therefore, researchers may consider using the MAQ and Flavell's perspective instead.

The ability to find differences between individuals' metacognitive abilities may be related to whether one is measuring metacognitive knowledge or measuring metacognitive control. Difference in metacognitive knowledge is measurable. Differences in metacognitive control, however, may be much harder to pinpoint (recall that Reder, 1987, 1996 insists that such processes are beyond conscious control). Therefore, in order to discriminate between individuals' at monitoring ability, we first need to determine if metacognitive control can be accurately measured.

Recent research has focused on the distinction between feeling-of-knowing judgments (FOK) and judgments-of-learning (JOL) (Plude, Nelson & Scholnick, 1998). A FOK asks whether you can information in the future. It is thought to represent a judgment of retrieval whether it is in long-term or short-term storage. On the other hand, a JOL is thought to reflect whether the information is learned or stored. Nelson claims JOLs are a short-term memory judgment (1996).

However, research indicates a JOL is more accurate when the judgment is delayed (Dunlosky & Nelson, 1992). Dunlosky and Nelson found more accurate judgments-of-learning in a paired-associate task when the stimulus and response were presented.
together, but not when only the stimulus was presented alone (1997). JOLs are also thought to be associated with strategy use for delayed judgments (Dunlosky & Nelson, 1994). For example, larger JOLs were found in a paired-associate task when an imagery, rather than rote memory, strategy was used. In the second part of the same study greater JOLs were reported for distributed practice over massed practice, when the judgment was delayed. It was hypothesized that immediate judgments reflect subjects' focusing on transient aspects of the task, leading to inaccurate predictions in immediate judgments. If judgments-of-learning need to be delayed to be accurate, this cast doubt on the JOL as being short-term memory judgment or encoding judgment.

Supporters of this distinction rebut by highlighting that each judgment is sensitive to different aspects of the task (Carroll, Nelson & Kirwan, 1997). In one study, subjects memorized various noun-noun lists in a paired-associate task. Some lists were composed of related nouns and others were comprised of unrelated noun-noun pairings. JOL and FOK judgments were made at 2 week and 6 week interval tests. Results indicated JOLs were more sensitive to lists when the lists they were composed of related material. FOK, on the other hand, were sensitive to overlearning. While these judgments may be related to different aspects of retrieval, they may not represent different aspects of learning.

While Nelson and Narens (1990) assert that monitoring judgment reflects how accessible a target is, there are an increasing number of studies opposing this view. Reder’s research has asserted that familiarity with the information is an important factor (1987; 1996; Reder & Ritter, 1992). Reder has focused on strategy selection in problem-solving and response latencies. Her conclusion is that FOKs represent an initial judgment of familiarity with the information. Metcalfe, Schwartz, and Jouquim (1993) agree with Reder when they compare judgments to the number of presentations of the to-be-learned
material. More accurate FOKs were related to the number of presentations of the material. Koriat (1993; 1994; 1995), in examining high FOK judgments with inaccurate responses, concluded that FOK judgments represented access to the material, not the actual retrieval of information. The work of Bjork and his colleagues (Benjamin, Bjork, & Schwartz, 1998) on the inferential aspects of FOK judgments adds to the consensus that confidence judgments are not reflective of direct access to target.

The paired-associate task in the current study with the interference component may certainly contribute to Reder’s work. Further use of this paradigm may contribute to the cue familiarity hypothesis. Observing subject attempts to acquire new information over several trials with an interference component may help determine whether the judgment is based on familiarity with the information or if the FOK is a measure of target availability. If it is a measure of availability there should be less interference between what is encoded and what is new, or potentially interfering, information.
APPENDIX I

INVENTORIES AND TESTS
METACOGNITIVE SCALES

KNOWLEDGE OF COGNITION

1. Declarative knowledge: knowledge about learning and one's cognitive skills and abilities

2. Procedural knowledge: knowledge about how to use strategies

3. Conditional knowledge: knowledge about when and why to use strategies

REGULATION OF COGNITION

1. Planning: planning, goal setting, and allocating resources

2. a) Organizing: implementing strategies and heuristics that help one manage information

   b) Information management: organizing, elaborating, summarizing, and selectively focusing on important information

3. Monitoring: on-line assessment of one's learning or strategy use

4. Debugging: strategies used to correct performance errors or assumptions about the task or strategy use

5. Evaluation: post-hoc analysis of performance and strategy effectiveness
QUESTIONS BY CATEGORY

METACOGNITIVE KNOWLEDGE

Declarative Knowledge. Items 5, 10, 12, 16, 17, 20, 32, 46 (8)

Procedural Knowledge. Items 3, 14, 27, 33 (4)

Conditional Knowledge. Items 15, 18, 29, 35 (5)

METACOGNITIVE REGULATION

Planning. Items 4, 6, 8, 22, 23, 42, 45 (7)

Strategy. Items 9, 13, 30, 31, 37, 39, 41, 43, 47, 48 (10)

Monitor. Items 1, 2, 11, 21, 28, 34, 49 (7)

Debug. Items 25, 40, 44, 51, 52 (5)

Evaluate. Items 7, 19, 24, 36, 38, 50 (6)
### METACOGNITIVE ASSESSMENT INVENTORY

We would like you to respond to the questions in this packet by indicating how true or false each statement is about you. If a statement is always true, write the number 5 in the blank provided to the right of each statement. Your responses are scored anonymously, so please answer as truthfully as you can.

<table>
<thead>
<tr>
<th>ALWAYS</th>
<th>SOMETIMES</th>
<th>SOMETIMES</th>
<th>ALWAYS</th>
</tr>
</thead>
<tbody>
<tr>
<td>FALSE</td>
<td>FALSE</td>
<td>NEUTRAL</td>
<td>TRUE</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

1. I ask myself periodically if I am meeting my goals.  
2. I consider several alternatives to a problem before I answer.  
3. I try to use strategies that have worked in the past.  
4. I pace myself while learning in order to have enough time.  
5. I understand my intellectual strengths and weaknesses.  
6. I think about what I really need to learn before I begin a task.  
7. I know how well I did once I finish a test.  
8. I set specific goals before I begin a task.  
9. I slow down when I encounter important information.  
10. I know what kind of information is most important to learn.  
11. I ask myself if I have considered all options when solving a problem.  
12. I am good at organizing information.  
13. I consciously focus my attention on important information.  
14. I have a specific purpose for each strategy I use.  
15. I learn best when I know something about the topic.  
16. I know what the teacher expects me to learn.
17. I am good at remembering information.
18. I use different learning strategies depending on the situation.
19. I ask myself if there was an easier way to do things after I finish a task.
20. I have control over how well I learn.
21. I periodically review to help me understand important relationships.
22. I ask myself questions about the material before I begin.
23. I think of several ways to solve a problem and choose the best one.
25. I ask others for help when I don't understand something.
26. I can motivate myself to learn when I need to.
27. I am aware of what strategies I use when I study.
28. I find myself analyzing the usefulness of strategies while I study.
29. I use my intellectual strengths to compensate for my weaknesses.
30. I focus on the meaning and significance of new information.
31. I create my own examples to make information more meaningful.
32. I am a good judge of how well I understand something.
33. I find myself using helpful learning strategies automatically.
34. I find myself pausing regularly to check my comprehension.
35. I know when each strategy I use will be most effective.
36. I ask myself how well I accomplished my goals once I'm finished.
37. I draw pictures or diagrams to help me understand while learning.
38. I ask myself if I have considered all options after I solve a problem.
39. I try to translate new information into my own words.
40. I change strategies when I fail to understand.
41. I use the organizational structure of the text to help me learn.

42. I read instructions carefully before I begin a task.

43. I ask myself if what I'm reading is related to what I already know.

44. I re-evaluate my assumptions when I get confused.

45. I organize my time to best accomplish my goals.

46. I learn more when I am interested in the topic.

47. I try to break studying down into smaller steps.

48. I focus on overall meaning rather than specifics.

49. I ask myself questions about how well I am doing while I am learning something new.

50. I ask myself if I learned as much as I could have once I finish a task.

51. I stop and go back over new information that is not clear.

52. I stop and reread when I get confused.
Please circle whether you agree (T) or disagree (F) with each statement.

T  F  1. I would prefer a task that is intellectual, difficult and important to one that is somewhat important but does not require much thought.

T  F  2. I tend to set goals that can be accomplished only by expending considerable mental effort.

T  F  3. Learning new ways to think doesn’t excite me very much.

T  F  4. I am hesitant about making important decisions after thinking about them.

T  F  5. I prefer just to let things happen rather than try to understand why they turned out that way.

T  F  6. I have difficulty thinking in new and unfamiliar situations.

T  F  7. The idea of relying on thought to make my way to the top does not appeal to me.

T  F  8. The notion of thinking abstractly is not appealing to me.

T  F  9. I often think as hard as I have to.

T  F  10. I don’t reason well under pressure.

T  F  11. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.

T  F  12. I find little satisfaction in deliberating hard and for long hours.

T  F  13. I more often talk with other people about the reasons for and possible solutions to international problems than about gossip of tidbits of what famous people are doing.

T  F  14. More often than not, more thinking just leads to more errors.

T  F  15. I don’t like to have the responsibility of handling a situation that requires a lot of thinking.

T  F  16. I appreciate opportunities to discover the strengths and weaknesses of my own reasoning.
T  F  17. I feel relief rather than satisfaction after completing a task that required a lot of mental effort.

T  F  18. Thinking is not my idea of fun.

T  F  19. I try to anticipate and avoid situations where there is a likely chance I will have to think in depth about something.

T  F  20. I prefer watching educational to entertainment programs.

T  F  21. I think best when those around me are very intelligent.

T  F  22. I prefer my life to be filled with puzzles that I must solve.

T  F  23. I would prefer complex to simple problems.

T  F  24. Simply knowing the answer rather than understanding the reasons for the answer to a problem is fine with me.

T  F  25. It's enough for me that something gets the job done. I don't care how or why it works.
Recall Task
General Knowledge Questions

Example 1. What is the name of the horse-like animal with black and white stripes?

Example 2. What is the last name of the author who wrote "The Old Man and the Sea"?

1. What is the last name of the British admiral who won the Battle of Trafalgar?
2. What is the last name of Dagwood's boss in the comic strip "Blondi"?
3. For which country is the Drachma the monetary unit?
4. What is the capital of Jamaica?
5. What is the name of the brightest star in the sky excluding the sun?
6. In what profession was Emmett Kelly?
7. What are people who make maps called?
8. What are people called who explore caves?
9. What was the name of King Arthur's sword?
10. What is the last name of the first American author to win the Nobel prize for literature?
11. What is the last name of the man who was most responsible for photographing the U.S. Civil War?
12. What is the last name of the songwriter who wrote the song "I Love Paris"?
13. What was the last name of the captain of the British ship "Bounty" when the mutiny occurred?
14. What was the last name of the female star of the movie "Casablanca"?
15. What is the name of the mountain range that separates Asia from Europe?
16. What is John Kenneth Galbraith's profession?
17. What is the capital of Finland?
18. Who was the leader of the Argonauts?

19. What is the last name of the woman who founded the American Red Cross?

20. What is the name of the extinct reptiles known as “Terrible Lizards”?

21. What is the name of the North Star?

22. What was the last name of Billy the Kid?

23. What was the last name of the boxer who was known as the “Manassa Mauler”?

24. What is the capital of New York?

25. What is the last name of the composer of the “Maple Leaf Rag”? 
## Answer Sheet
### Recall Task

<table>
<thead>
<tr>
<th>Confidence Interval</th>
<th>Example 1.</th>
<th>Example 2.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 20 40 60 80 100</td>
<td>Example 1.</td>
<td>Example 2.</td>
</tr>
<tr>
<td>0 20 40 60 80 100</td>
<td>1.</td>
<td>2.</td>
</tr>
<tr>
<td>0 20 40 60 80 100</td>
<td>3.</td>
<td>4.</td>
</tr>
<tr>
<td>0 20 40 60 80 100</td>
<td>5.</td>
<td>6.</td>
</tr>
<tr>
<td>0 20 40 60 80 100</td>
<td>7.</td>
<td>8.</td>
</tr>
<tr>
<td>0 20 40 60 80 100</td>
<td>9.</td>
<td>10.</td>
</tr>
<tr>
<td>0 20 40 60 80 100</td>
<td>11.</td>
<td>12.</td>
</tr>
<tr>
<td>0 20 40 60 80 100</td>
<td>13.</td>
<td>14.</td>
</tr>
<tr>
<td>0 20 40 60 80 100</td>
<td>15.</td>
<td>16.</td>
</tr>
<tr>
<td>0 20 40 60 80 100</td>
<td>17.</td>
<td>18.</td>
</tr>
<tr>
<td>0 20 40 60 80 100</td>
<td>19.</td>
<td>20.</td>
</tr>
<tr>
<td>0 20 40 60 80 100</td>
<td>21.</td>
<td>22.</td>
</tr>
<tr>
<td>0 20 40 60 80 100</td>
<td>23.</td>
<td>24.</td>
</tr>
<tr>
<td>0 20 40 60 80 100</td>
<td>25.</td>
<td></td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Recognition Task
General Knowledge Questions

Example 1. What is the name of the horse-like animal with black and white stripes?
   a. Tiger       d. Mule
   b. Zebra       e. Gnu
   c. Giraffe

Example 2. What is the last name of the author who wrote “The Old Man and the Sea”? 
   a. Steinbeck   d. London
   b. Fitzgerald  e. Hemingway
   c. Faulkner

1. What is the last name of the British admiral who won the Battle of Trafalgar?
   a. Nelson      d. Powell
   b. Napoleon    e. Montgomery
   c. MacArthur

2. What is the last name of Dagwood’s boss in the comic strip “Blondi”? 
   a. Smith       d. Bowen
   b. Dithers     e. Swetnam
   c. Hill

3. For which country is the Drachma the monetary unit?
   a. India       d. Greece
   b. Iran        e. Israel
   c. Turkey

4. What is the capitol of Jamaica?
   a. Kingston    d. Montego
   b. Havana     e. Barbados
   c. Georgetown

5. What is the name of the brightest star in the sky excluding the sun?
   a. Polaris     d. Pleiades
   b. Rigel       e. Sirius
   c. Andromeda

6. In what profession was Emmett Kelly?
   a. Dancer       d. Clown
   b. Artist      e. Diplomat
   c. Archeologist

7. What are people who make maps called?
   a. Cartographers d. Calligraphers
   b. Ontologists   e. Ethnographers
   c. Choreographers
8. What are people called who explore caves?
   a. Cave expedition   d. Photographer
   b. Engineer         e. Archeologist
   c. Spelunkers

9. What was the name of King Arthur’s sword?
   a. Excalibur        d. Miriam
   b. Galahad          e. Thanatos
   c. Ivanhoe

10. What is the last name of the first American author to win the Nobel prize for literature?
    a. Mitchner        d. Emerson
    b. Thoreau         e. Henry
    c. Faulkner

11. What is the last name of the man who was most responsible for photographing the U.S. Civil War?
    a. Weston           d. Kodak
    b. Carter           e. Daguerre
    c. Brady

12. What is the last name of the songwriter who wrote the song “I Love Paris”?
    a. Berlin           d. Hammerstein
    b. Porter           e. Gershwin
    c. Sullivan

13. What was the last name of the captain of the British ship “Bounty” when the mutiny occurred?
    a. Bligh           d. Cook
    b. Ahab            e. Queeg
    c. Hook

14. What was the last name of the female star of the movie “Casablanca”?
    a. Bacall           d. Young
    b. Hayward          e. Bergman
    c. Hepburn

15. What is the name of the mountain range that separates Asia from Europe?
    a. Salt Range       d. Pyrenees
    b. Ural             e. Sudety
    c. Appalachian

16. What is John Kenneth Galbraith’s profession?
    a. Photographer     d. Economist
    b. Architect        e. Poet
    c. Biographer
17. What is the capital of Finland?
   a. Copenhagen        d. Prague
   b. Helsinki          e. Stockholm
   c. Oslo

18. Who was the leader of the Argonauts?
   a. Hercules          d. Jason
   b. Pericles          e. Ulysses
   c. Achilles

19. What is the last name of the woman who founded the American Red Cross?
   a. Anthony          d. Anthony
   b. Curie            e. Barton
   c. Nightingale

20. What is the name of the extinct reptiles known as “Terrible Lizards”?
   a. Dinosaurs        d. Archaeopteryx
   b. Amphibians       e. Trilobites
   c. Jurassic

21. What is the name of the North Star?
   a. Sirius           d. Betelgeuse
   b. Polaris          e. Antares
   c. Vega

22. What was the last name of Billy the Kid?
   a. Bonney           d. Cole
   b. Blane            e. Hickock
   c. James

23. What was the last name of the boxer who was known as the “Manassa Mauler”?
   a. Marciano         d. Willard
   b. Clay             e. Sullivan
   c. Dempsey

24. What is the capital of New York?
   a. Rochester        d. Buffalo
   b. Ithaca           e. Syracuse
   c. Albany

25. What is the last name of the composer of the “Maple Leaf Rag”?
   a. Joplin           d. Hammerstein
   b. Basie            e. Armstrong
   c. Sousa
### Confidence Interval

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>20</td>
<td>40</td>
<td>60</td>
<td>80</td>
<td>100</td>
</tr>
</tbody>
</table>

Example 1. A) B) C) D) E)

Example 2. A) B) C) D) E)

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>20</td>
<td>40</td>
<td>60</td>
<td>80</td>
<td>100</td>
</tr>
</tbody>
</table>

1. A) B) C) D) E)

2. A) B) C) D) E)

3. A) B) C) D) E)

4. A) B) C) D) E)

5. A) B) C) D) E)

6. A) B) C) D) E)

7. A) B) C) D) E)

8. A) B) C) D) E)

9. A) B) C) D) E)

10. A) B) C) D) E)

11. A) B) C) D) E)

12. A) B) C) D) E)

13. A) B) C) D) E)


15. A) B) C) D) E)

16. A) B) C) D) E)

17. A) B) C) D) E)

18. A) B) C) D) E)

19. A) B) C) D) E)

20. A) B) C) D) E)

21. A) B) C) D) E)

22. A) B) C) D) E)

23. A) B) C) D) E)


25. A) B) C) D) E)
## Paired-Associate Task Trial 1

(AB)

<table>
<thead>
<tr>
<th>Learn</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banana 17</td>
<td>Orange 23</td>
</tr>
<tr>
<td>Grapefruit 13</td>
<td>Mushroom 10</td>
</tr>
<tr>
<td>Potato 26</td>
<td>Pear 16</td>
</tr>
<tr>
<td>Turnip 15</td>
<td>Turnip 15</td>
</tr>
<tr>
<td>Orange 23</td>
<td>Lemon 06</td>
</tr>
<tr>
<td>Cucumber 01</td>
<td>Corn 14</td>
</tr>
<tr>
<td>Bean 07</td>
<td>Tangerine 30</td>
</tr>
<tr>
<td>Asparagus 25</td>
<td>Spinach 22</td>
</tr>
<tr>
<td>Carrot 27</td>
<td>Asparagus 25</td>
</tr>
<tr>
<td>Lemon 06</td>
<td>Potato 26</td>
</tr>
<tr>
<td>Mushroom 10</td>
<td>Banana 17</td>
</tr>
<tr>
<td>Pineapple 05</td>
<td>Cherry 20</td>
</tr>
<tr>
<td>Spinach 22</td>
<td>Kiwi 08</td>
</tr>
<tr>
<td>Kiwi 08</td>
<td>Bean 07</td>
</tr>
<tr>
<td>Corn 14</td>
<td>Pineapple 05</td>
</tr>
<tr>
<td>Artichoke 03</td>
<td>Cucumber 01</td>
</tr>
<tr>
<td>Pear 16</td>
<td>Radish 11</td>
</tr>
<tr>
<td>Radish 11</td>
<td>Carrot 27</td>
</tr>
<tr>
<td>Cherry 20</td>
<td>Grapefruit 13</td>
</tr>
<tr>
<td>Tangerine 30</td>
<td>Artichoke 03</td>
</tr>
</tbody>
</table>
Paired-Associate Task Trial 2

(Abr)

<table>
<thead>
<tr>
<th>Learn</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bean 07</td>
<td>Pineapple (05)</td>
</tr>
<tr>
<td>Kiwi 08</td>
<td>Asparagus (25)</td>
</tr>
<tr>
<td>Pear 16</td>
<td>Artichoke (03)</td>
</tr>
<tr>
<td>Tangerine 30</td>
<td>Radish (11)</td>
</tr>
<tr>
<td>Cucumber 01</td>
<td>Lemon (06)</td>
</tr>
<tr>
<td>Lemon 06</td>
<td>Cherry (20)</td>
</tr>
<tr>
<td>Spinach 22</td>
<td>Potato (26)</td>
</tr>
<tr>
<td>Radish 11</td>
<td>Cucumber (01)</td>
</tr>
<tr>
<td>Corn 14</td>
<td>Banana (17)</td>
</tr>
<tr>
<td>Pineapple 05</td>
<td>Kiwi (08)</td>
</tr>
<tr>
<td>Orange 23</td>
<td>Mushroom (10)</td>
</tr>
<tr>
<td>Banana 17</td>
<td>Spinach (22)</td>
</tr>
<tr>
<td>Artichoke 03</td>
<td>Bean (07)</td>
</tr>
<tr>
<td>Cherry 20</td>
<td>Tangerine (30)</td>
</tr>
<tr>
<td>Grapefruit 13</td>
<td>Turnip (15)</td>
</tr>
<tr>
<td>Carrot 27</td>
<td>Grapefruit (13)</td>
</tr>
<tr>
<td>Potato 26</td>
<td>Corn (14)</td>
</tr>
<tr>
<td>Mushroom 10</td>
<td>Pear (16)</td>
</tr>
<tr>
<td>Turnip 15</td>
<td>Orange (23)</td>
</tr>
<tr>
<td>Asparagus 25</td>
<td>Carrot (27)</td>
</tr>
</tbody>
</table>
Paired-Associate Task Trial 3
(AC)

<table>
<thead>
<tr>
<th>Learn</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grapefruit</td>
<td>Corn (05)</td>
</tr>
<tr>
<td>Carrot</td>
<td>Kiwi (21)</td>
</tr>
<tr>
<td>Orange</td>
<td>Bean (14)</td>
</tr>
<tr>
<td>Radish</td>
<td>Spinach (15)</td>
</tr>
<tr>
<td>Turnip</td>
<td>Tangerine (30)</td>
</tr>
<tr>
<td>Asparagus</td>
<td>Cherry (20)</td>
</tr>
<tr>
<td>Pineapple</td>
<td>Cucumber (01)</td>
</tr>
<tr>
<td>Kiwi</td>
<td>Radish (11)</td>
</tr>
<tr>
<td>Pear</td>
<td>Potato (18)</td>
</tr>
<tr>
<td>Banana</td>
<td>Artichoke (03)</td>
</tr>
<tr>
<td>Corn</td>
<td>Turnip (06)</td>
</tr>
<tr>
<td>Lemon</td>
<td>Grapefruit (12)</td>
</tr>
<tr>
<td>Cherry</td>
<td>Orange (23)</td>
</tr>
<tr>
<td>Potato</td>
<td>Asparagus (07)</td>
</tr>
<tr>
<td>Spinach</td>
<td>Mushroom (10)</td>
</tr>
<tr>
<td>Tangerine</td>
<td>Pineapple (22)</td>
</tr>
<tr>
<td>Artichoke</td>
<td>Lemon (25)</td>
</tr>
<tr>
<td>Bean</td>
<td>Banana (04)</td>
</tr>
<tr>
<td>Cucumber</td>
<td>Pear (24)</td>
</tr>
<tr>
<td>Mushroom</td>
<td>Carrot (09)</td>
</tr>
</tbody>
</table>
Paired-Associate Task

Answer Sheet

Confidence Interval

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>80</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>8.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>9.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>10.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>11.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>12.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>13.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>14.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>15.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>16.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>17.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>18.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>19.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>20.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
</tbody>
</table>
APPENDIX II

TABLES AND FIGURES REFERENCED IN TEXT
Table 1

Means and Standard Deviations of Groups on MAI Factor 2

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Monitor</td>
<td>146.67</td>
<td>8.89</td>
</tr>
<tr>
<td>Mid Monitor</td>
<td>125.46</td>
<td>3.96</td>
</tr>
<tr>
<td>Low Monitor</td>
<td>107.96</td>
<td>5.47</td>
</tr>
</tbody>
</table>
Table 2

Bias Score Means on Feeling-of-knowing Task

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Monitor</td>
<td>0.09</td>
<td>0.17</td>
</tr>
<tr>
<td>Med Monitor</td>
<td>0.11</td>
<td>0.20</td>
</tr>
<tr>
<td>Low Monitor</td>
<td>0.05</td>
<td>0.15</td>
</tr>
</tbody>
</table>
Table 3

Bias Means on Paired-associates Task

<table>
<thead>
<tr>
<th>Group</th>
<th>AB</th>
<th>SD</th>
<th>ABr</th>
<th>SD</th>
<th>AC</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>0.03</td>
<td>0.13</td>
<td>0.06</td>
<td>0.11</td>
<td>0.05</td>
<td>0.13</td>
</tr>
<tr>
<td>Medium</td>
<td>0.10</td>
<td>0.18</td>
<td>0.04</td>
<td>0.19</td>
<td>0.06</td>
<td>0.20</td>
</tr>
<tr>
<td>Low</td>
<td>0.05</td>
<td>0.15</td>
<td>0.04</td>
<td>0.11</td>
<td>0.03</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table 4

Calibration Means on Feeling-of-knowing Task

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Monitor</td>
<td>0.46</td>
<td>0.55</td>
</tr>
<tr>
<td>Med Monitor</td>
<td>0.51</td>
<td>0.55</td>
</tr>
<tr>
<td>Low Monitor</td>
<td>0.50</td>
<td>0.55</td>
</tr>
</tbody>
</table>
Table 5

**Calibration Means on Paired-associates Task**

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Monitor</td>
<td>1.24</td>
<td>0.88</td>
</tr>
<tr>
<td>Medium Monitor</td>
<td>1.34</td>
<td>0.80</td>
</tr>
<tr>
<td>Low Monitor</td>
<td>1.48</td>
<td>1.20</td>
</tr>
<tr>
<td>ABr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Monitor</td>
<td>1.49</td>
<td>0.62</td>
</tr>
<tr>
<td>Medium Monitor</td>
<td>1.22</td>
<td>0.70</td>
</tr>
<tr>
<td>Low Monitor</td>
<td>1.49</td>
<td>0.78</td>
</tr>
<tr>
<td>AC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Monitor</td>
<td>1.67</td>
<td>0.75</td>
</tr>
<tr>
<td>Medium Monitor</td>
<td>1.26</td>
<td>1.15</td>
</tr>
<tr>
<td>Low Monitor</td>
<td>1.19</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Table 6

Means and Standard Deviations of Groups on MAI Factor 1

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>High monitor</td>
<td>71.81</td>
<td>4.60</td>
</tr>
<tr>
<td>Mid monitor</td>
<td>63.14</td>
<td>1.39</td>
</tr>
<tr>
<td>Low monitor</td>
<td>54.23</td>
<td>5.97</td>
</tr>
</tbody>
</table>
Table 7

Means and Standard Deviations of Groups on Total MAI Score

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>High monitor</td>
<td>214.60</td>
<td>10.06</td>
</tr>
<tr>
<td>Mid monitor</td>
<td>189.04</td>
<td>6.22</td>
</tr>
<tr>
<td>Low monitor</td>
<td>164.48</td>
<td>10.22</td>
</tr>
</tbody>
</table>
Table 8

Correlation Matrix Between MAI and NFC

<table>
<thead>
<tr>
<th></th>
<th>MAI Factor 1</th>
<th>MAI Factor 2</th>
<th>NFC</th>
<th>Persist</th>
<th>Conf</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor 1</td>
<td>0.78*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor 2</td>
<td>0.94*</td>
<td>0.53*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFC</td>
<td>0.55*</td>
<td>0.47*</td>
<td>0.49*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persist</td>
<td>0.42*</td>
<td>0.35*</td>
<td>0.38*</td>
<td>0.86*</td>
<td></td>
</tr>
<tr>
<td>Confid</td>
<td>0.51*</td>
<td>0.51*</td>
<td>0.42*</td>
<td>0.80*</td>
<td>0.53*</td>
</tr>
<tr>
<td>Complex</td>
<td>0.48*</td>
<td>0.38*</td>
<td>0.45*</td>
<td>0.88*</td>
<td>0.62*</td>
</tr>
</tbody>
</table>

* p < 0.01
Figure 1

Bias Means by Ability Group

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 2

Calibration Means by Ability Group

- High monitors
- Mid monitors
- Low monitors

Calibration scores

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 3

Paired-associates Bias Means
Figure 4

Paired-associates Calibration Means
REFERENCES


Monitoring and control processes (pp. 139-159). New Jersey: Lawrence Erlbaum Associates.


(Eds.) **Metacognition: Knowing about knowing** (pp. 48-70). Cambridge, Mass.: MIT Press.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


VITA

Graduate College
University of Nevada, Las Vegas

Brett Douglas Campbell

Local Address:
1675 Silver Glen Ave
Las Vegas, NV 89123

Degrees:
Bachelor of Arts, Psychology, 1990
Weber State University

Thesis Title: An Evaluation Between Confidence Judgments and Differences in Monitoring Ability

Thesis Examination Committee:
Chairperson, Dr. Alice J. Corkill, Ph.D.
Committee Member, Dr. Gregory Schraw
Committee Member, Dr. Frank Dempster
Graduate Faculty Representative, Dr. Mark Gaudagnoli