4-29-2008

The Relationship between Executive Functions and Metacognitive Strategy Learning and Application

Eva Mashock Jansiewicz
Georgia State University

THE RELATIONSHIP BETWEEN EXECUTIVE FUNCTIONS AND METACOGNITIVE STRATEGY LEARNING AND APPLICATION

by

Eva Jansiewicz

Under the Direction of Dr. Robin Morris

ABSTRACT

This project examined whether the executive functions of set maintenance and switching, as assessed by neuropsychological testing, were predictors of set maintenance and switching within a more ecologically valid task that used metacognitive strategies during reading comprehension tasks as a framework for evaluation. Gaze times to key words during reading were used as an indirect measure of strategy use. A few significant relationships were found between set maintenance and set switching on the neuropsychological measures and the strategy learning and application tasks. Participants were more likely to switch to appropriate strategies in a situation in which they were given free choice of strategies to use, and in which characteristics of the text pulled for the use of a particular strategy. In contrast, participants were less consistent with expected strategy use when they had just learned a strategy and were asked explicitly to apply it to a text that did not pull for use of a particular strategy. Factors such as visual scanning, motor speed, working memory, and passage comprehension affected the relationship between executive functions and the more ecologically valid task.

INDEX WORDS: Executive functions, Metacognitive strategies, Reading comprehension, Set Maintenance, Set Switching

THE RELATIONSHIP BETWEEN EXECUTIVE FUNCTIONS AND METACOGNITIVE STRATEGY LEARNING AND APPLICATION

by

EVA MASHOCK JANSIEWICZ

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2008

THE RELATIONSHIP BETWEEN EXECUTIVE FUNCTIONS AND METACOGNITIVE STRATEGY LEARNING AND APPLICATION

by

EVA MASHOCK JANSIEWICZ

Committee Chair: Robin Morris

Committee: Mary Morris

Erin McClure

David Washburn

Electric Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2008

DEDICATION

To my daughter Alexis Sofia. May you find joy in your life.

ACKNOWLEDGMENTS

This work could not have been completed without the help of my mentor, Robin Morris. He has been a mentor in the truest sense, helping me to find my way in my career and the rest of my life. I have been inspired by his incredible talent and empathy. Thank you, Robin, for being willing to support me during these life-changing years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

When individuals seek a neuropsychological evaluation, they are often interested in learning about their strengths and weaknesses and in receiving recommendations for assistance in treating their problems. However, very little research supports whether such recommendations are appropriate, based on cognitive strengths and weaknesses of the individual. There is a clear need for increased knowledge about how neuropsychological assessment findings can be used for effective treatment planning. Specifically, we have much more to learn about how the information gathered during neuropsychological assessment can predict patients' behaviors in the real world, but most importantly, in treatment.

This project attempted to determine the extent to which executive functions, as assessed by neuropsychological testing, can be used as predictors of a person's ability to apply different metacognitive strategies during reading a more real world type of task. Specifically, this project will examine if the executive functions of set maintenance and switching are significantly related to set maintenance and switching within the context of a more ecologically valid task which required using different metacognitive strategies during reading comprehension tasks.

The Relationship Between Executive Functions and Metacognitive Strategies

Executive functions are an important element of what makes us adaptable, thinking human beings. They include various aspects of organizing, monitoring, and directing mental

processes. Examples of areas included in the realm of executive functions include self-awareness and monitoring of behavior and its consequences, maintenance of a course of action, switching a course of action when appropriate, problem-solving, concept formation and abstract reasoning. Neuropsychological testing assesses these functions though tasks such as card sorting, category formation, mazes, verbal reasoning, reading tasks, tower construction, go-no-go tasks, and connecting numbers and letters or dots in certain patterns. Such tasks are often multi-faceted, testing several aspects of executive function simultaneously. Partitioning the different aspects involved can be difficult, but an analytic approach is likely to be the most appropriate means of describing the executive processes (Anderson, 2002).

When brain injury occurs, executive functions are often impaired as a result. Deficits in executive functions can have a crucial role in determining the extent of social and vocational outcome after brain injury (Sohlberg & Mateer, 1989). Fortunately, research has shown that executive functions can be remediated to some extent through cognitive interventions. Based on a review of the literature on cognitive remediation, the Brain Injury Interdisciplinary Special Interest Group of the American Congress of Rehabilitation Medicine recommends the training of formal problem solving strategies and their application to everyday situations for individuals with stroke or traumatic brain injury (TBI). It is this focus on everyday, real world outcomes that drives one of the main goals of this study: whether neuropsychological texts can predict such real world functioning?

Metacognitive strategies are techniques to increase an individual's awareness of their thought processes and actions while completing tasks. In this sense, they can be seen as the

behavioral output of executive functions (i.e. the executive function of working memory is seen in metacognitive terms as individuals being able to monitor how well they are keeping information in memory). Barkley (1996) acknowledges that both the terms executive function and metacognition suffer from an ambiguity of definition. Many describe metacognition as "thinking about thinking." One of the earliest definitions of metacognition was "knowledge and cognition about cognitive phenomena" (Flavell, 1979).

Metacognitive strategies can be taught and their effective use has been shown to improve task performance. Research by Lodico (1983) attempted to teach children the value of strategies and to monitor the relationship between strategy use and task performance. Children who were taught to monitor strategy effectiveness recognized that their better performance on tasks was due to effective strategy use, chose the more effective strategy when given a choice, and explained that their choice was made because they believed it would improve their performance. However, children who did not receive monitoring training could not explain why they had chosen particular strategies or gave explanations that were unrelated to their performance (Lodico, 1983).

Further investigation of both metacognition and executive function could be beneficial in several ways. First, it would help to link the concepts of executive function in neuropsychology and metacognition in educational intervention, so that professionals could better understand each others' language. Second, an examination of the relationship between executive functions and metacognition skills could help to predict which individuals with executive function problems might benefit from strategy instruction. For example, if it is found that individuals with low

levels of executive function also have little strategy use and do not benefit from strategy instruction, it would be better to provide compensatory strategies (such as multimodal presentation) or more structured teaching rather than conceptual strategy instruction during remediation.

Although the concepts of executive functions and metacognitive strategies have many similarities, they typically differ in their ecological validity. Neuropsychological tests, though, have been found to have a moderate ability to predict everyday behaviors in neurologically typical populations (Sbordone & Guilmette, 1999). There has also been limited research on the relationship between knowledge of metacognitive strategies and use of these strategies, but the correlation may be fairly high because strategies are often trained through practicing the strategy, and then individuals are asked to utilize the same or different strategy in a similar situation as an outcome measure. Executive function tasks and metacognitive strategy tasks also have different content. Executive function tests often measure a narrow sample of behavior in response to novel stimuli (i.e. naming the colors of words) and generalize these findings to the wider construct of executive function. On the other hand, metacognitive strategy assessments measure a sample of strategy utilization behavior, which is the same as the outcome measure.

Strategy Use During Reading Comprehension as a More Ecologically-Valid Measure of Executive Functioning

Metacognitive skills are crucial in reading comprehension. In order to understand text, readers must be able to monitor their comprehension and apply strategies to improve their

comprehension if they do not understand. Trabasso and Bouchard (2002) define comprehension strategies as specific procedures that are learned and allow for active, goal-centered, self-regulated reading. Metacognitive reading comprehension strategies are most often taught and modeled by the teacher and then practiced by the students alone or in groups. The common element to all these teaching methods is that they encourage the student to think about what they are reading and take an active role in the process, maintaining and switching strategies during reading. This process clearly requires higher order cognitive functioning and will be the focus of the current study.

Unfortunately, much of past reading comprehension strategy research has relied upon self-report measures as a measurement approach. In this approach, the reader tells the experimenter which strategies they were using and what they were thinking during different parts of the reading experiment. The strength of this method is that the subjective experience of the reader can be accessed, but the drawback is lack of an objective measurement of the actual strategy being used, or how different changing strategies might be applied. Also, self-report measures often have a low correlation with observed behavior, especially for individuals with neurological impairment (Burke, Smith & Imhoff, 1989). More objective measures of strategy use would help to increase the validity of such paradigms. One such objective measure that might show promise is eye tracking. Although the research literature does not show any examples of the use of eye tracking technology to study metacognitive strategy use, its ability to provide good experimental precision, objective measurement, and ability to observe natural reading, including regressions, make this a potential new method for studying reading comprehension strategy use if an appropriate paradigm can be developed.

The Potential of Using Eye Tracking to Study Metacognitive Strategy Use and Changes During Reading

The introduction of eye tracking technology has revolutionized the study of reading. Using this technology, an individual can read a text while external eye tracking equipment monitors the location and duration of their gaze at different parts of the text, and a report of these variables can be generated for analysis. Although much debate remains about whether eye movements are directly correlated with cognitive processes during reading (the eye-mind problem), most research supports the idea that eye movements are linked to language related variables, but that not all higher level comprehension processes are completed within a single fixation (Rayner & Carroll, 1984). However, there is evidence from error recovery research (in which the reader needs to fix a previous misinterpretation of text later in the text) that some comprehension processes are completed immediately upon the first fixation of a word (Just & Carpenter, 1984). Other processes occur at the ends of sentences and are called wrap-up effects (Just & Carpenter, 1984).

When readers move their eyes forward in the text, they generally fixate on the very next word. They skip over a word only 93% of the time (Just & Carpenter, 1984). Content words are fixated 83% of the time, while function words are fixated 38% of the time. Word length and frequency are estimated to account for 69% of mean gaze duration on words (Just & Carpenter, 1984). In this study, the semantic role of the word in a sentence will be used as a key marker word to study strategy use.

Qualities of the text can have considerable impact on eye movements while reading. Lower-level characteristics that influence gaze time include length and frequency of the word, its semantic role in the sentence, if it is a novel word, and if it introduces a new topic (Just & Carpenter, 1984).

As text becomes more difficult, most readers slow down. Their average saccade length decreases, average fixation duration increases, and the frequency of regressions increase. There is also increased fixation on content words in difficult texts (Just & Carpenter, 1984). In cases in which a word is misspelled, fixations on critical words are increased (Zola, 1982). Also, words unexpected within the context of a passage have longer fixations than other words in the text (Ehrlich & Rayner, 1981). Misleading, or garden path sentences have also been used to show that when readers are expecting one word, but see a contextually unexpected one, they will spend extra time looking at the text in order to comprehend the information presented (Frazier & Rayner, 1982). Similarly, it has been found that if a word is difficult to process, the reader may make several different fixations upon it and that those fixations may be longer in duration than for other words (Just & Carpenter, 1984). Based on this literature, Just and Carpenter (1984) use gaze duration, or the sum of all fixations on a word as a measure for analyses. They have found that this measure has a precise quantitative relationship to the processes they hypothesize occur during reading (Just & Carpenter, 1984). In this framework, the assumption is that changes in gaze time represent changes in the text, or its semantic requirements. This inferred relationship will underlie that paradigm designed for this study.

Directions given to the reader can also have an effect on the way that text is processed. Just and Carpenter (1984) have found through their research that instructions to monitor for possible inconsistencies will elicit many regressions in readers, but instructions not to reread will elicit very few regressions. Also, characteristics of the reader can have a significant effect on text processing. For example, working memory span can predict reading comprehension. Traditional tests of short-term memory are not correlated with reading comprehension performance, but tests of reading span are (Just & Carpenter, 1984).

Anaphors are pronouns or other words that refer to another word or concept, and their effects on eye monitoring patterns have been well-studied. Anaphors are constructed in the following manner. In the following sentences, “Bill spent the day talking to Pam. He realized that he really liked her.”, the word “her” refers to Pam, and the word “he” refers to Bill. Both “he” and “her” are anaphors in the second sentence, referring to the people mentioned in the previous sentence. Such anaphors are called reference anaphors, and are personal or demonstrative pronouns that refer to objects that were previously mentioned (Yuill & Oakhill, 1988). Yuill and Oakhill (1988) describe three other types of anaphors. The first is ellipsis or substitution, which indicates the replacement of one item for another. In substitution, the replacement word is marked lexically, as in “Sam took a test. Jane did too.” in which “did” replaces “took a test.” In ellipsis, the same principle applies, but the replacement is not marked, as in “Are you taking a test? Yes, I am.” in which “I am” means “I am taking a test.” The final type, lexical cohesion, is a semantic link in the text, such as in “I went to the pool and jumped off the diving board.” The anaphor allows the assumption that the diving board is part of the pool.

In a study of anaphors, Carpenter and Just (1978) found that readers tend to make regressive fixations to the referent of the pronoun, and that if the referent was ambiguous (2 or more nouns could qualify as the referent), readers looked back to the referent 50% of the time, and increased their gaze time to the referent. These results suggest that ambiguous anaphors can increase gaze time to their referents. Similarly, Ehrlich and Rayner (1983) found that subjects regressed to antecedents more frequently when there was an inconsistent antecedent to a pronoun (i.e. inconsistent gender). They concluded that subjects regress to antecedents when instructions or task demands encourage this behavior. Ehrlich and Rayner (1983) also found that when the distance between the pronoun and referent was greater, shorter saccades and longer fixation durations were present, presumably to allow extra time to resolve the anaphor. Based on this research literature, one could instruct reading comprehension strategies focused on anaphor analysis while using gaze time for those anaphors as an indirect measure of the strategy being used.

Current Study

The current study's primary goal was to determine if there was a significant relationship between certain elements of executive function tests and those same elements in a situation in which metacognitive strategy use and application are required. Although there are many different components of executive functions related to metacognitive strategy use, the core elements of interest are the ability to determine an appropriate strategy in a specific situation, switch strategies to use the most appropriate strategy, and maintain an appropriate strategy until a new strategy would be more effective or is required based on instructions. Thus, in the current

project, set maintenance and set switching were examined using tests of executive function, and these results were compared to measures of set maintenance and set switching within the context of a more ecologically valid task requiring the use of different metacognitive strategies during a reading comprehension paradigm. This study was not designed to be a study of reading comprehension or metacognitive strategy use, but instead a study of the *relationship* between set maintenance and set switching as measured by executive function tests and set maintenance and set switching as measured with a more ecologically valid task that used different metacognitive strategies during a reading comprehension task.

To assess maintaining and switching between metacognitive strategies during reading in a more objective manner, an eye tracking device will be used during the reading of texts to measure gaze time. The paradigm will determine the location and duration of the participant's gaze to certain key words in the text that will be considered indicators of the strategies being used. The paradigm to be used involves teaching the use of two different metacognitive strategies that are easily objectified: anaphor re-reading (measured by the amount of time looking at key words that help to resolve anaphors) and determining the main idea (measured by the amount of time spent looking at key main idea words in the texts). The use of these two strategies (anaphor, main idea) was a proxy through which to investigate set maintenance and set switching within a more ecologically valid context.

In the current study, each strategy was specifically trained before the participant was asked to apply it. The anaphor strategy was taught using the methods of Yuill and Oakhill (1988). They conducted a pre-training session, in which participants were read a script

describing the different anaphors in sentences. They went through pre-training examples until they could identify the referent of each anaphor without error. In the current study, the same method was used for the training tasks, except that participants read the script themselves, as they were college students. The main idea strategy was trained using the methods of Mullen (1987) instructing readers to determine which points in a paragraph are narrow or global in scope. Mullen (1987) trained participants that if there is one global point, this is the main idea, or if there is more general point, the relationship between the points will be the main idea. She also indicated that the main idea may be implied and is often found in the first two sentences of a paragraph. Participants were presented with examples in which the main idea was located in different parts of a paragraph, and then asked to complete several examples and identify the main idea in each paragraph. After receiving each strategy training (main idea and anaphor), all participants were asked to complete a series of up to 8 training trials in which they were asked to apply the trained strategy to a text. These training trials were completed to assure that each participant was able to use the strategy correctly before proceeding with the eye tracking task. Participants were provided corrective feedback for all incorrect answers to the sample questions. Training trial scores were calculated for each of the training trial sets for a maximum of eight per strategy training. The texts for the training tasks and training trials can be found in Appendices A and B.

During the study, participants completed eye tracking during a series of reading comprehension tasks to assess their strategy maintenance and strategy switching as required by the study design. The eye tracking reading comprehension tasks contained several texts (blocks), each of which required similar or different strategy use when compared to the preceding block.

Figure 1 shows the progression of these blocks during the experiment. The block labels are located on the left side of the figure, and blue arrows representing strategy switches and purple arrows representing strategy maintenance are located on the right side of the figure. First, participants completed the baseline text block. Then they received training on the first (main idea) strategy and completed the training trials. Next, they completed the second text block, being told to use the main idea strategy just taught on the next two paragraphs. How similar their strategy use was between the 1st and 2nd paragraphs was considered a measure of strategy maintenance. Then they were trained on the second (anaphor) strategy and were required to use it on the next two paragraphs. Moving from Block two (main idea) to Block 3 (anaphor) required a switch in strategies. Participants then completed the fourth and fifth blocks, being told to use the strategy that works best. Strategy use was calculated by summing eye fixation times (gaze times) to key words considered to indicate strategy use. Strategy training was presented in a counterbalanced order during the course of the experiment, with half of the participants receiving training on the main idea strategy first, and half receiving training on the anaphor strategy first. However, the results will be presented with the main idea strategy results in the second block and the anaphor strategy results in the third block for conceptual clarity.

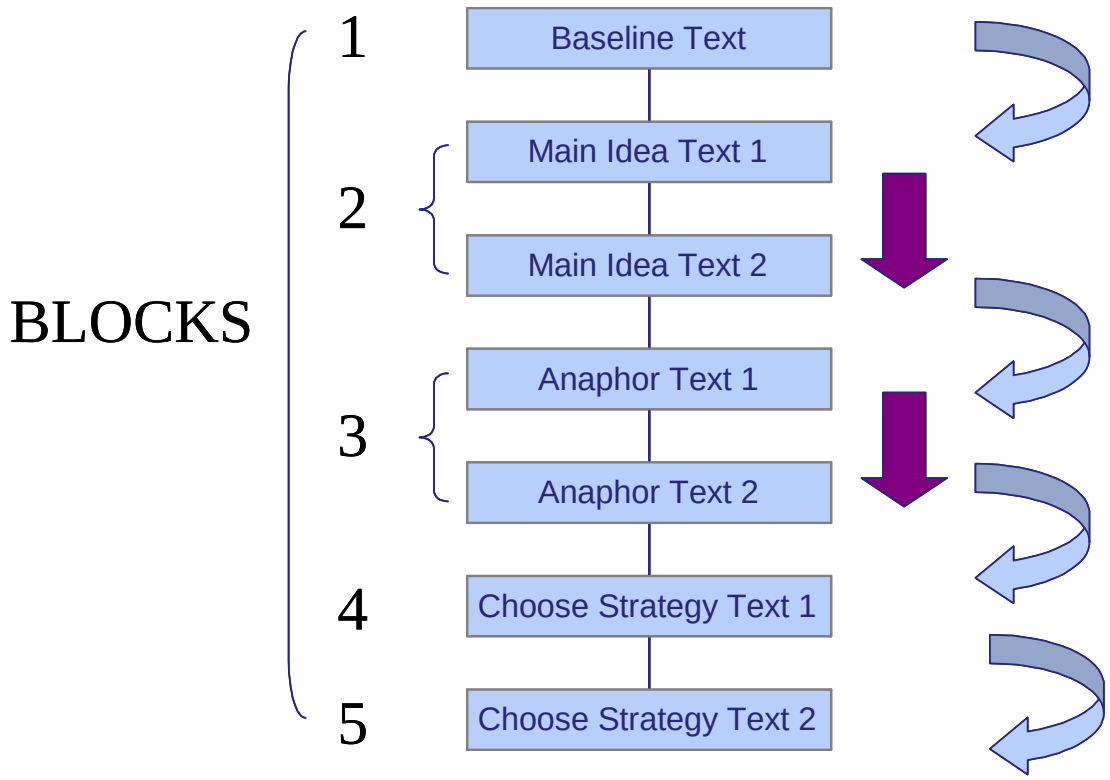

Figure 1.
Progression of Eye Tracking Portion of Experiment

After each text, participants completed cloze comprehension questions about the texts they just read and were asked to circle main idea and anaphor words in printed versions of the texts as a reliability check. Next, they completed a self-report measure of what strategies they used at different points in the experiment and how useful they found them (Strategy Use Survey) while continuing to view the printed texts for reference purposes. Next, they completed the Metacognitive Reading Strategies Questionnaire (MRSQ), a standardized measure of metacognitive strategy use during reading comprehension tasks.

Finally, participants completed several neuropsychological assessment measures in the following order: WAIS-III Digit Span, SB-IV Memory for Sentences, WJ-III Passage

Comprehension, DKEFS Trail Making Test, DKEFS Color Word Test, BCT, and CPT. Some of these measures were used to assess the executive function constructs of set maintenance and set switching and vigilance (attention) and some were used as covariates to assess working memory, reading comprehension, motor speed, and visual scanning.

Research Hypotheses

It was hypothesized that set maintenance and set switching, as assessed by executive function measures, would be positively correlated with set maintenance and set switching within the context of a more ecologically valid measure of metacognitive strategy maintenance or switching in a reading comprehension paradigm. Five specific hypotheses were made in this study:

Hypothesis 1: It was predicted that scores on set maintenance as assessed by executive function measures would be significantly negatively correlated with the number of training trials necessary to reach learning criterion for each new strategy taught.

Hypothesis 2: It was predicted that set maintenance as assessed by executive function tests, and vigilance, as assessed by the attention task, would be significantly and positively correlated with set maintenance as assessed within the reading comprehension paradigm. Similarly, set switching as assessed by executive function tests would be significantly and positively correlated with set switching as assessed within the reading comprehension paradigm.

Hypothesis 3: It was predicted that participants would respond to experimenter instructions and use the main idea strategy more than the anaphor strategy in the second block while using the anaphor strategy more than the main idea strategy in the third block of the experimental paradigm. They were also expected to rate these specific strategies as more used on the strategy survey. It was further predicted that participants would use specific strategies more consistent with qualities of the text rather than strategies inconsistent with the qualities of the text in the fourth and fifth blocks of the experiment. Participants were expected to use the main idea strategy more and rate it as more useful for the fourth block than for the fifth block and use the anaphor strategy more and rate it as more useful for the fifth block than for the fourth block.

Hypothesis 4: It was predicted that scores on set switching and maintenance on neuropsychological tests and from the eye tracking data would be significantly positively correlated with participants' scores on a measure of self-reported use of reading comprehension strategies (Metacognitive Reading Strategies Questionnaire (MRSQ; Taraban, Kerr, & Rynearson, 2004).

Hypothesis 5: It was predicted that strategy use, as assessed by eye tracking data, would be significantly positively correlated with the participants' subjective judgments of when they used each strategy at each stage of the strategy use tasks (Strategy Use Survey).

Chapter 2

Methods

Participants

A series of pilot studies was conducted prior to initiating the main study. Appendix G describes these pilot studies.

For the main study, thirty-eight undergraduate students were recruited from a large public university in the Southeastern United States. Participants were recruited from a subject pool of introductory psychology class students wishing to receive class credits for participating in research studies. The population at this university is diverse in terms of race, ethnicity, and age. Many individuals at this university are able to acquire a full-tuition scholarship through the state, so there is also a great deal of socioeconomic diversity. The effect sizes of previous studies in this area (Trabasso & Bouchard, 2002) have been in the medium range (.50), according to Cohen's guidelines (1977). Thus, this study sought an effect size of .50. A power analysis (PS Power, Dupont & Plummer, 1997) determined that a sample size of 33 at a power level of .80 would be adequate for paired t-test analyses of average fixation times on the eye tracking tasks which formed the basis of the analysis.

Four of the 38 participants were excluded from the final analyses. One participant had red-green color blindness, which invalidated his results on the DKEFS Color-Word Test. Three other participants had gaze time values of 0 milliseconds for at least one text in the eye tracking

data. These values were considered very unlikely and an artifact of poor calibration, so the participants were excluded from the final sample.

The final analyses included 34 participants: 25 female, and nine male. Self-reported ethnicities were: 47% White, 26% Black, 12% Asian, 6% Hispanic/Latino, 6% Biracial/Bicultural, and 3% Other Race. 26 participants were right-handed, 7 were left-handed, and one did not report their handedness. Participants ranged in age from 18 to 67 years of age, and the mean age was 24.18 (10.22). Ninety-one percent of participants were between 18 and 27 years of age, and three participants who were older: 44, 51, and 67 years of age, respectively. College GPA ranged from 1.55-3.81 on a 4-point scale, with a mean of 2.95 (.47).

Materials

The materials selected were used to assess two constructs of executive function (set switching and maintenance) and the application of these two constructs to a reading comprehension task in college students.

Standardized Measures

The Rosenbaum Pocket Vision Screener was used to tests participants' vision to assure that they could see the materials presented in the experiment. All participants' vision, as tested with the Rosenbaum Pocket Vision Screener, was in the normal range, except for poor vision in

one eye for two participants. All participants were able to read and respond to directions presented on the screen on which the eye tracking texts appeared.

The executive functions of set maintenance and set switching were assessed using the Booklet Category Test, 2nd edition (BCT) (DeFilippis & McCampbell, 1997), and the Trail-Making Test (TMT) and Color-Word Interference Test (CWIT) from the Delis-Kaplan Executive Function System (DKEFS), (Delis, D.C, Kaplan, E., & Kramer, J.H., 2001). The BCT, TMT, and CWIT were each used assess set maintenance as well as set switching.

The BCT has a long history of validity and reliability studies and is considered to be moderate to high in both (Lezak, Howieson, & Loring, 2004). The average adult scoring in the 50th percentile has 29-37 errors on the BCT (Heaton et al., 2004), so participants were unlikely to have a ceiling effect on this measure. For the BCT, without explicit instructions, a participant must determine, understand, and apply a problem-solving rule throughout a subtest, based only on the feedback they are given whether they are right or wrong. They must then determine a new rule each time the rule changes. In subtests one through five, the rule is different in each subtest. The rule in subtest six is the same as in subtest five. Subtest seven is a memory test of the previous items.

Kucera-Thompson (2003) derived three factors from the BCT, based on results from item/sequence analysis on the BCT, such as the number of items in a run and items representing a change in strategy. The factors were forming mental set (Power), switching mental set (Total Errors, Flexibility), and maintaining mental set (Speed, Maintain). Kucera-Thompson's (2003)

factors were used as guidelines to calculate set switching and set maintenance scores for the current study. In the current study, set maintenance for the BCT was calculated as the maximum span of items correct for each subtest, with the sum of all subtests comprising the set maintenance index score. Set switching was calculated by determining whether the individual had 10 items in a row correct on each subtest. If they did, their set switching score was 1 (yes), if not, their set switching score was 0 (no). The set switching scores from each subtest were summed to create a set switching index score. As subtest six has the same rule as subtest five, switching was not calculated on this subtest, only maintenance. Subtest seven is a memory test and was not included in the calculations of set switching and maintenance.

The DKEFS has test-retest correlations of .43 to .73 and moderate internal consistency reliability coefficients (Lezak, Howieson, & Loring, 2004). Raw scores on the TMT and CWIT are recorded in seconds to completion, so ceilings are rarely possible on these measures. For the TMT, Conditions 2 and 3, respectively, involve the timed connection of a series of numbers and letters, which selectively assesses set maintenance and automaticity since there is only one task requirement that needs to be maintained. However, Condition 4 requires the participant to switch quickly between connecting numbers and letters, and is more sensitive to cognitive inflexibility, or switching (Lezak, Howieson, & Loring, 2004). Conditions 1 and 5 of the TMT, respectively, assess scanning and motor speed, which can be used as covariates. On the CWIT, Conditions 1 and 2 measure the participant's ability to maintain a set by rapidly naming colors or reading words. However, in Condition 4, the participant needs to switch quickly between two naming strategies (naming colors and reading words). Thus, Conditions 1 and 2 on this measure were used to assess set maintenance, and Condition 4 was used to assess set switching abilities.

Condition 3 is an interference condition in which the participant is to name color of the ink of inconsistently colored words, and could be considered a more complex and difficult set-maintenance task.

As sustained attention is closely linked with the maintenance component of executive functions (Barkley, 1996; Morris, 1996), the Continuous Performance Test (CPT; Conners, 1992), which lasts 14 minutes, was used. During this measure, a series of letters appears one at a time on the screen. The participant presses a button every time a letter other than X appears. The CPT normative sample includes data from adults with brain disorders and individuals with ADHD (Lezak, Howieson, & Loring, 2004). Errors of omission were analyzed to assess set maintenance, as omission is associated with lack of vigilance, which is associated with the ability to maintain set.

Working memory was assessed in this study, as it is thought to be a central component in executive function (Diamond, 1997) and may be a necessary prerequisite to the cognitive operations involved in set maintenance and set switching. In this study, two tasks were used to assess working memory. Both were used as covariates in the analyses. The first task was Digit Span from the Weschler Adult Intelligence Scale – Third Edition (WAIS-III) (Wechsler, 1997) which requires the participant to repeat a string of numbers forwards and backwards. The WAIS-III has both high reliability (.96 test-retest for the full-scale IQ) and high content, criterion, and construct validity. The second task was Sentence Memory from the Stanford-Binet – Fifth Edition (SB-V) (Thorndike et al., 1986) which was also used to assess word span. In this task, the participant repeats orally presented sentences of increasing complexity. The SB-V has

reliability and validity in the range of .80 to .90. Word span has been shown to be highly correlated with reading comprehension (Just & Carpenter, 1984).

Reading comprehension was assessed in the study as a covariate to explore the possibility that it might affect the relationship between set maintenance and set switching as measured by executive function tasks and the strategy use tasks among participants. The Passage Comprehension subtest from the Woodcock-Johnson Tests of Achievement – III (Woodcock, McGrew, & Mather, 2001) was used to assess reading comprehension. The WJ-III tests are frequently used and have both high reliability and high validity.

This study also used the Metacognitive Reading Strategies Questionnaire (MRSQ; Taraban, Rynearson, & Kerr, 2004) to assess participants' amount of metacognitive strategy use during everyday reading of their college textbooks. The MRSQ is a self-report measure of reading strategies created through summarizing the major reading comprehension strategies in 38 published reports of strategies used by adult skilled readers and taught at the elementary and secondary levels. Taraban et al. (2000) found that their self-report measure of reading strategy use given to college students reliably discriminated between participants with higher and lower GPAs. Strategy ratings also discriminated between participants with higher and lower Reading and English ACT scores. A principal components analysis of the scores of 575 college students who completed the measure revealed two constructs: analytic cognitions supporting reading comprehension and pragmatic behaviors assessing studying and academic performance. Cross-validation was completed with a second sample, and revealed high validity and internal consistency (Cronbach's alpha = .82).

Experimental Measures

In order to measure metacognitive strategy use, a procedure for exposing participants to strategies and assessing metacognitive strategy was implemented. Participants were taught two different metacognitive reading comprehension strategies (main idea, anaphor) and then their inferred strategy use was assessed using eye tracking gaze time to key words while they read texts on a computer screen. The gaze time for key words linked to each strategy within the texts was recorded as an index of the participants' reading strategy. An underlying assumption of this approach was that gaze time to key words consistent with the required strategy use would be longer than gaze times to key words not consistent with the strategy. Participants were taught two different metacognitive strategies – 1) attending to main ideas (main idea strategy) and 2) rereading text to clarify information (anaphor strategy). In order to assess their use of the main idea strategy, the amount of time participants gazed at main idea words in the texts was calculated. For example, for a passage about dinosaurs, the word dinosaur was considered a main idea word. In order to assess their use of anaphor rereading of texts to clarify information, anaphors were introduced in the text and the amount of time spent looking at key words (anaphor referents) that clarifed the anaphors was calculated. For example, the following sentences might be presented: "John went to the beach. He had a good day." In this series of sentences, "John" would be the key word that would clarify the anaphor "he." In addition, a randomly-selected set of nonkey words, typically nouns, were also analyzed to evaluate the specificity of the strategy effects vs. those nonkey words not directly relevant to either strategy (control words).

The texts used for this study were modified from training texts in a college level textbook on reading comprehension strategies (Smith, 2005). Modifications that were made included shortening the texts so that they would fit on two screens during computer presentation, adjusting texts so each contained 8 main idea words, 8 anaphor words, and 8 control words (nouns not related to strategy use) and including two inconsistent (i.e. incorrect gender) anaphors for each text. Anaphors and their referents were always included together on a single screen of text. The text was centered vertically and horizontally on the screen and was in Courier New font, 18 point size. The monitor on which the text was presented was 36 cm wide and 28 ½ cm high. Participants' eyes were located 66 cm from the monitor while reading the texts. The presentation of a text in EPrime automatically triggered the recording of eye movements in DQW. These texts are presented in Appendix C.

The eye tracking task was divided into five blocks. Prior to completing each new strategy, participants completed the following three tasks in sequence: eye calibration, instructions for the task, and silent reading of the texts while eye movements were being recorded. Between blocks, participants received training on the specific reading comprehension strategies. The baseline (first) block preceded training on any strategy use, then the training on the main idea strategy took place, followed by the second block of texts, then the training on the second strategy (anaphor) took place, followed by the third block of texts. The second and third blocks used texts that were structured in such a way that both strategies were equally useful in reading the texts. This was done by having the main idea embedded in the middle or end of each text, rather than in the beginning, where it would be easier to identify. Also, each of the texts in the second and third blocks had the same number (two) of inconsistent (i.e. incorrect gender)

anaphors. For texts in the second and third blocks, participants were instructed to use the strategy just taught. In contrast, in the fourth and fifth blocks, the texts that were used were structured so that one of the strategies taught was more useful than the other strategy in reading the text. The text for Block 4 pulled for use of the anaphor strategy. It was constructed so that the main idea of the text was in the first sentence and clearly stated. This manipulation was expected to make the main idea strategy less necessary for understanding this text, so that the anaphor strategy would be used instead. On the other hand, the text for block five pulled for use of the main idea strategy. In this text, the main idea was located in the middle and end of the text. However, the anaphors for this text were simpler than in other texts and were all gender specific, such as “he” and “him”, rather than “it” or “one” as in other texts. Also, there were fewer anaphor referents in this text which were repeated multiple times, instead of having one anaphor per referent, as in other texts. Due to these modifications, participants were expected to use the main idea strategy more for this text, because the anaphors were fairly simple and did not require much anaphor strategy use. For the texts in the fourth and fifth blocks, participants were not asked to apply a particular strategy, but instead to use whichever of the two strategies worked best.

Thus, in the second and third blocks, participants were asked to use a particular strategy, and in the fourth and fifth blocks they were asked to pick whichever of the two strategies worked best. This experimental manipulation allowed for the possibility of a double dissociation between task demands and strategy use to be investigated, to see whether participants were more likely to use certain strategies when they were more effective, but not specifically instructed.

Printed versions of the texts were used as a reliability check on the participants' ability to identify main idea and anaphor words in the texts. These texts are located in Appendix C. On this measure, participants circled all the main idea and anaphor words they could find. Also, participants completed a self-assessment in which they were asked how much they used the two strategies while completing the eye tracking texts and how useful they found the strategies while reading the eye tracking texts. This measure is located in Appendix D. Participants also completed a questionnaire (Metacognitive Reading Strategies Questionnaire) assessing how much they use metacognitive strategies for reading comprehension in their college classes. This measure is located in Appendix E.

Finally, comprehension questions using cloze technique (fill in the blank with the word that makes sense based on the passage) were also given after the eye tracking reading task to increase motivation and to assure that the participants were paying attention to the text and comprehending the material. This measure is located in Appendix F. One question was given per text. If participants answered 7/7 (100%) of the questions correctly, they proceeded to the next measure in the experiment. However, if they received a score of less than 7/7 correct on the questions, they received an additional set of 7 sample questions to complete. Whether or not this second set of sample questions was completed at 100% accuracy, participants proceeded to the next experimental measure after completion. Participants were provided corrective feedback for all incorrect answers to the questions. Participants who completed the first seven questions correctly were given automatic credit for the second set of questions and were given a score of 14/14. Participants who did not have a score of 100% on the first set of questions and received

the second set of sample questions received a score of the number correct out of the 14 questions they completed.

Apparatus

Figure 2 shows the setup of the apparatus used. The texts were presented on a Dell monitor that was 36 cm wide and 28 ½ cm high. The monitor was raised to allow an appropriate eye angle when reading the texts. The bottom of the monitor was located 21 cm above the table on which it rested, and 17 ½ cm back from the front edge of the table. Participants sat at a second table, the far edge of which was located 44 ½ cm from the edge of the table that the monitor rested on. A chin rest was attached to the table at which the participants sat, and the participants' eyes were located 38 ½ cm from the far edge of the table when they placed their chins in the chin rest. Thus, the participants' eyes were located 66 cm from the Dell monitor. The chin rest was usually set at the highest level (37 cm), but was adjusted for certain participants when a good eye image could not be obtained at this height. Participants sat in an adjustable chair that allowed them to adjust the height of their bodies to the chin rest.

Eye tracking was conducted using a head mounted eye imaging system (Eyecam SN 01-501-0622 from IScan Inc., Boston, MA). A Panasonic Video monitor was used to monitor the scene as viewed by the camera on the eye imaging system, and a Sony Trinitron monitor was used to monitor the eye movements in real time. Two desktop computers were also used during the eye tracking, one to present the stimuli to the participant in EPrime 1.0 (Psychology Software

Tools, Inc., 2001), and the other for the experimenter to calibrate and monitor the eye movements of the participant. The software program DQW Version 1.0 (ISCAN, Inc., 1997) was used to collect eye movement data. Figure 2 shows the setup of the eye tracking equipment.

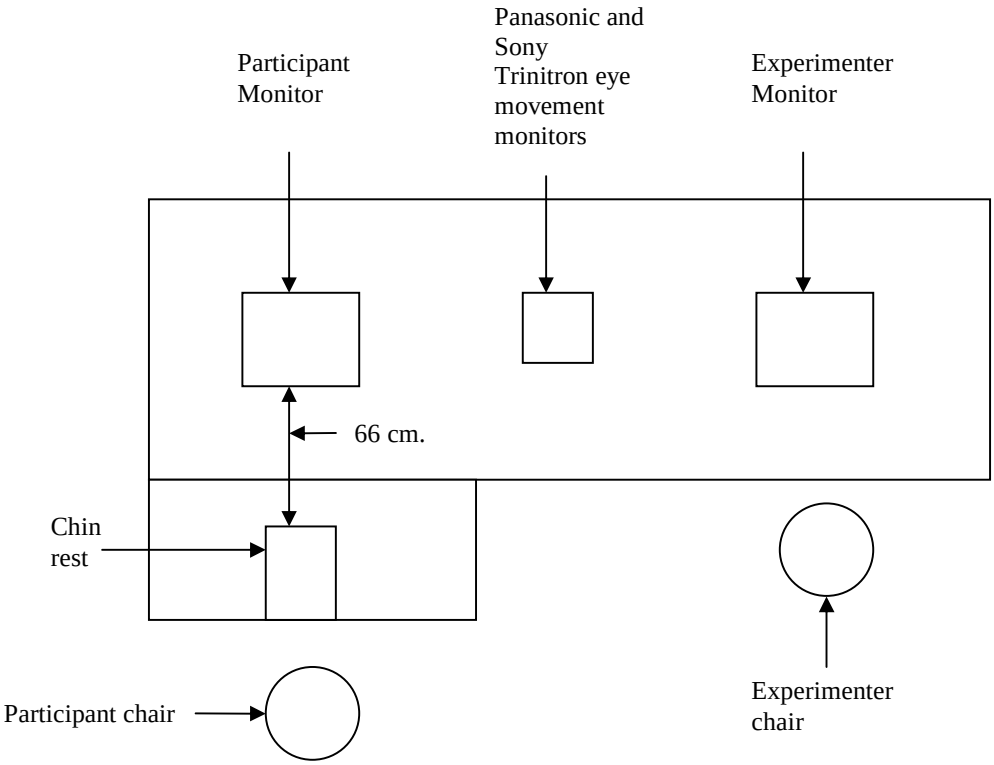

Figure 2.
Setup of Eye Tracking Equipment

The locations of main idea and anaphor key words and the control words on the screen were measured in centimeters from the top and left side of the screen while the text was displayed in Eprime. These locations were converted into the pixel coordinates of the scene monitor. Data from the DQW software program were recorded in pixels ranging from 45-488

pixels horizontally and 47-450 pixels vertically. This was the same range as in the scene monitor, allowing for the appropriate conversion of the location of key words in centimeters to pixel coordinates. Data from the DQW program were converted into Excel files for data analysis. An Excel spreadsheet was created which summed the fixation times for all main idea and anaphor key words and control words in each text in seconds.

A nine-point calibration was completed with participants before any eye tracking data were collected. Additionally, this calibration was conducted at the beginning of each block of the eye tracking task to check that the eye tracker had an accurate measure of where each participants' eyes were looking on the computer monitor. The nine point calibration screen included nine crosshair fixation points evenly spaced at the top left, center left, top right, center left, center, center right, bottom left, bottom center, and bottom right areas of the screen. The calibration points covered the full viewing area of the screen. Participants were asked to look at each crosshair in turn, and the experimenter calibrated each point as the participant looked at it. Next, the experimenter asked the participant to again look at each point in turn and the experimenter checked the calibration. If proper calibration was not achieved, the sequence was repeated until proper calibration was achieved.

Procedure

The experiment proceeded in a predetermined order. First, informed consent was completed with participants and any questions participants had were answered. If the participants

agreed to continue in the experiment, they completed a Participant Background Form which asked for basic demographic information. Next, they completed a basic vision screening with the Rosenbaum Pocket Vision Screener.

Next, they completed the eye tracking reading comprehension task. First, they completed the baseline block. Then they received training on the first strategy and completed sample questions. Next, they completed the second block, being told to use the strategy just taught. Then they were trained on the second strategy and completed sample questions. Next, they completed the third block, being told to use the strategy just taught. Then they completed the fourth and fifth blocks, being told to use the strategy that works best.

Participants then completed the cloze comprehension questions and were asked to circle main idea and anaphor words in the printed texts. Next, they completed the self-report measure of what strategies they used at different points in the experiment and how useful they found those strategies (Strategy Use Survey) while continuing to view the printed texts for reference purposes. Next, they completed the Metacognitive Reading Strategies Questionnaire (MRSQ).

Finally, participants completed the neuropsychological assessment measures in the following order: WAIS-III Digit Span, SB-IV Memory for Sentences, WJ-III Passage Comprehension, DKEFS Trail Making Test, DKEFS Color Word Test, BCT, and CPT.

Data Analysis

Strategy Training Orders

This study used two different strategy training orders (1 and 2) to control for order and fatigue effects. Table 1 shows the different trials and strategy instructions across the experiment for counterbalance orders 1 and 2. The texts used are provided in Appendix C. The texts used were presented in the same order for Orders 1 and 2, except for the last two texts, which were switched in Order 2. Data analysis focused initially on whether these order effects yielded different results, and if not, the orders were collapsed.

Table 1.
Strategy Instructions for Eye Tracking Trials by Counterbalance Order

Block	Text #	Order 1 (n = 17)	Order 2 (n = 17)
1		Baseline	Baseline
2	1	Main Idea Strategy	Anaphor Strategy
2	2	Main Idea Strategy	Anaphor Strategy
3	1	Anaphor Strategy	Main Idea Strategy
3	2	Anaphor Strategy	Main Idea Strategy
4		Choose Strategy	Choose Strategy
5		Choose Strategy	Choose Strategy

Eye Movement Data Preparation

There were two high value outliers (19689 and 22745 msec of gaze time) for two participants on one text passage. These values were interpreted as possible values of individuals who were spending a lot of time trying to use the required strategy. The range of values for the rest of the original passages was 634 - 9435 msec of gaze time. Data analyses were run with the original outlier values included and again with substituted data values which were one point greater than the largest non-outlier value in the distribution (9436 msec) in order to maintain their relationship to the other data. The results of the two sets of analyses did not differ in any significant way. The substituted data values were therefore kept in the final analyses in order to reduce the possible effects of these two outlier values on skewness. There were other outlier values at the low end of the distribution, but these were kept in the distribution because there were several of these values and they were spread across participants and texts. The summed gaze time values by participant and text number are included in Appendix H.

In order to obtain measures of maintenance and set switching related to strategy-specific use and increase the reliability of the eye tracking data, two ratios of eye tracking data were calculated and used as the basis of the analyses that examined the research hypotheses. A maintenance ratio was calculated by using a reference point for each individual's gaze times in an attempt to standardize gaze times within each person by dividing each individual's summed gaze time to strategy consistent key words by their summed gaze time to control words. The maintenance ratio was calculated using the following equation to assess strategy maintenance between texts 2 and 3 within Block 2 (main idea strategy):

$$\frac{\left|\text{Strategy Consistent Gaze Time for Text 2} - \text{Strategy Consistent Gaze Time for Text 3}\right|}{\left|\text{Control Words Gaze Time for Text 2} - \text{Control Words Gaze Time for Text 3}\right|}$$

This ratio took the absolute value of the difference of strategy consistent words (main idea) between texts 2 and 3. It also took the absolute value of the difference of control words between texts 2 and 3. Then the two absolute values were divided. Absolute values were used in order to capture the difference between the two texts without regard to direction of the difference. A lower value on this ratio indicated smaller differences in strategy consistent gaze time between the two texts, which was considered an indication of better maintenance of the strategy over time. A higher value on this ratio indicated that there was a larger difference in strategy consistent gaze time between the two texts, and thus less consistent maintenance.

A very similar equation was used to calculate the maintenance ratio between texts 4 and 5 (Block 3):

$$\frac{\left|\text{Strategy Consistent Gaze Time for Text 4} - \text{Strategy Consistent Gaze Time for Text 5}\right|}{\left|\text{Control Words Gaze Time for Text 4} - \text{Control Words Gaze Time for Text 5}\right|}$$

The values used for determining this ratio and the computed ratios per participant are shown in Appendix I. The mean ratios computed per block are also included in Appendix I. In analyses using the maintenance ratio, both the maintenance ratios for Blocks 2 (main idea) and 3 (anaphor) were used and were entered as two separate variables in each analysis.

A switch ratio was also computed between texts for each participant in order to obtain a measure of each participant's gaze time to words specifically associated with the new strategy they were expected to use compared to their previous strategy. This ratio divided summed gaze time to strategy-consistent words by gaze time to strategy-inconsistent words (i.e. anaphor words when instructed to apply the main idea strategy) for each block. The following formula was used:

$$\frac{\text{Strategy Consistent Gaze Time}}{\text{Strategy Inconsistent Gaze Time}}$$

This formula was calculated for each participant for each block of the study, except for the baseline block, in which no strategy was expected to be used. A higher value on the switch ratio indicated longer gaze time to the key words focused on the strategy that the participant was instructed to use. A lower value on the switch ratio indicated shorter gaze time to the strategy the participant was expected to use compared to gaze time to those key words important for the other (inconsistent) strategy. The switch ratios for all participants for all blocks are shown in Appendix I.

The switch ratios constructed were specific to particular blocks within the study. In order to determine if a significant switch had occurred between two blocks of the study, a comparison between the switch ratios between each pair of texts was made. If the switch ratio value for each of the two texts that were being compared had a value of one or above, it was considered a valid switch between the texts and coded with a value of 1. If the switch ratio value for either of the two texts that were being compared had a value below one, it was not considered a valid switch

between the texts and was coded with a value of 0. With this system, each participant received a score of 1 for each valid switch between texts and a score of 0 for all other comparisons between the texts. This system was used to ensure that for each switch, each participant was starting by using the correct strategy in the first text, and then correctly switching to using the other strategy in the second text. There were more consistent set switches between blocks 4 and 5 than between other blocks, so the switch value of 1 or 0 between blocks 4 and 5 was used as the measure of set switching on the eye tracking task for all analyses that follow.

Text Circling Task

The text circling task used the same texts as the eye tracking text. Participants were asked to identify (via circling) the main idea and anaphor words on paper. This text circling task was designed to provide another measure of the difficulty of each text and how well participants were complying with the strategy use instructions. A Repeated Measures MANOVA was used to determine whether there were differences in the number of anaphor and main idea key words circled in each text, and whether there were differences in the number of words circled between counterbalance orders 1 and 2.

Participants were only asked to circle both main idea and anaphor words for the last four texts, so eight variables were entered into the analysis – the number of main idea and anaphor key words circled for each of the four texts. For the first three texts, participants were asked to circle anaphor key words, main idea key words, or the words they thought they spent the most

time looking at, but not both anaphor and main idea word within the same text. Thus, these texts were not analyzed because they were not an appropriate comparison for the other texts.

There was a maximum score of eight main idea and eight anaphor key words that could be circled for each text. The results show an overall effect of different numbers of main idea words and anaphor words circled in the different texts, [$F(7,26) = 17.22$, $p<.05$]. Post-hoc pairwise comparisons of the eight variables (with Bonferroni corrections) showed that significantly fewer main idea key words were circled for Text 4 than for the other three texts, and that significantly more anaphor key words were circled for Text 6 than the other three texts (except Text 3). There were no other significant differences in the number of main idea and anaphor key words circled between the other texts. It appears that the accurate identification of main idea and anaphor words in most texts was similar. For the between-subjects portion of the analysis, there was not a significant difference between the number of main idea and anaphor key words circled for counterbalance orders 1 and 2 on the texts, [$F(2,32) = .18$, $p = .67$].

Combining Eye Tracking Data

Correlations were used to assess the level of consistency between the two trials within each block in the eye tracking tasks in which participants were instructed to use the main idea strategy and between the two trials within each block in which participants were instructed to use the anaphor strategy. This was done in order to determine whether these two sets of texts within blocks could be combined for the data analyses. The gaze time to key words for the strategy that

was targeted in each block (main idea or anaphor) was summed and these summed values were correlated using Pearson correlations. There were significant correlations between gaze time for the two main idea strategy trials (r=.35, p<.05) and gaze time for the two anaphor strategy trials (r=.51, p<.05). Since there were significant correlations between the two trials of each condition and since combining trails would improve the reliability of the results, the trial sets within a block were collapsed by averaging the gaze time for the two trials. Table 2 shows the order of the texts used in the final analyses for orders 1 and 2. The remainder of the analyses refer to these five blocks (1, 2, 3, 4 & 5).

Table 2.
Strategy Instructions of Eye Tracking Trials by Counterbalance Order – Combined Trials

Block	Order 1	Order 2
1	Baseline	Baseline
2	Main Idea Strategy	Anaphor Strategy
3	Anaphor Strategy	Main Idea Strategy
4 (Main Idea)	Choose Strategy	Choose Strategy
5 (Anaphor)	Choose Strategy	Choose Strategy

In several of the analyses that follow, the strategy used (main idea or anaphor) is the variable of interest. Orders 1 and 2 had different orders of these strategy use conditions. Thus, in order to simplify the analyses, the trial numbers from order 1 (shown in Table 3) will be used to refer to the following strategy use conditions for the remainder of the analyses.

Table 3.
Block Numbers for Strategy Use Conditions

Block	Strategy Use Condition
1	Baseline
2	Main Idea Strategy
3	Anaphor Strategy
4	Choose Strategy*
5	Choose Strategy*

* The Block 4 text “pulled” for main idea strategy use and the Block 5 text “pulled” for anaphor strategy use.

Chapter 3

Results

Descriptive Statistics of Dependent Variables

First, descriptive statistics on the executive function variables were computed. All scores were within the average range. These results are shown in Tables 4 and 5.

Table 4.
Descriptive Statistics of Executive Function Set Maintenance Measures

Variable	Mean	SD
Trail Making Test 2 Scale Score	11.35	1.77
Trail Making Test 3 Scale Score	12.23	1.87
Color Word Interference Test 1 Scale Score	10.26	2.45
Color Word Interference Test 2 Scale Score	11.53	2.21
Booklet Category Test Maintenance (sum of # in longest runs)	95.12	29.54
Booklet Category Test T Score	41.59	12.19
Continuous Performance Test % Omissions T Score	53.14	28.35

N=34

Table 5.
Descriptive Statistics of Executive Function Set Switching Measures

Variable	Mean	SD
Trail Making Test 4 Scale Score	11.35	1.77
Color Word Interference Test 4 Scale Score	10.26	2.45
Booklet Category Test Switching (number of correct switches out of 4)	3.20	0.81

N=34

Descriptive statistics of the planned covariate measures are shown in Table 6. These values were also in the average range.

Table 6.
Descriptive Statistics of Planned Covariate Measures

Variable	Mean	SD
Digit Span Scale Score	10.23	2.70
Sentence Memory T Score	47.44	8.48
Passage Comprehension Scale Score	103.20	9.07
Trail Making Test 1 Scale Score	11.18	1.99
Color Word Interference Test 3 Scale Score	10.26	2.72

N=34

Descriptive statistics on the self report and strategy training measures are shown in Table 7.

Table 7.
Descriptive Statistics of Self-Report and Strategy Training Measures

Variable	Mean	SD
MRSQ Analytic Index (out of 80)	55.82	9.22
MRSQ Pragmatic Index (out of 30)	20.00	4.87
Anaphor Training Trials (out of 8)	5.85	2.18
Main Idea Training Trials (out of 8)	7.50	1.05
Strategy Training Comprehension Questions (out of 14)	12.62	1.82

N=34

Intercorrelations between the measures of executive function and between the planned covariates were also computed. Tables 8 and 9 show the results of these analyses.

Table 8.
Intercorrelations of Executive Function Measures

	TMT 2	TMT3	TMT4	CWIT1	CWIT2	CWIT4	BCT	CPT
TMT 2		.71**	.46**	.07	.17	.22	.04	.33
TMT 3			.38*	-.04	.20	.22	.08	.21
TMT 4				.21	.45**	.42*	-.01	.17
CWIT 1					.63**	.60**	-.01	.08
CWIT 2						.66**	.01	.12
CWIT 4							.24	.04
BCT								.41*

TMT 2, 3, & 4 = Trail Making Tests 2, 3, & 4; CWIT 1, 2, & 4 = Color-Word Interference Tests 2, 3 & 4; BCT = Booklet Category Test Maintenance Score; CPT = Continuous Performance Test % Omissions; * = p<.05; ** = p<.01

Table 9.
Intercorrelations of Planned Covariate Measures

	TMT 1	TMT 5	CWIT 3	Digit Span	Sent Mem	Pass Comp
TMT 1		-.09	.07	.05	.18	-.24
TMT 5			.16	.12	-.13	.07
CWIT 3				.01	.10	.23
Digit Span					.57**	.17
Sent Mem						.48**

TMT 1 & 5 = Trail Making Tests 1 & 5; CWIT 3 = Color-Word Interference Test 3; Digit Span = Digit Span from WAIS-III; Sent Mem = Sentence Memory from SB-IV; Pass Comp = Passage Comprehension from WJ-III; ** = p<.01

Order Effects

A Repeated Measures MANOVA was used to assess order effects for the eye tracking data. Gaze time to main idea words and anaphor words for each of the five texts in the eye tracking task was compared for the two orders, 1 and 2. The between-subjects portion of this analysis showed that the counterbalance orders 1 and 2 were not significantly different from each other [$F(2,32)=.56$, $p=.46$]. Thus, the data from the two counterbalance orders were combined.

Evaluating the Validity of the Eye Movement Data

Before further analyses of the eye movement data were completed, four validity analyses were conducted in order to determine whether participants responded to the eye movement task as expected.

The first validity analysis examined the possibility that there was systematic change in reading time across texts that was not specific to strategy use instructions. This was examined by comparing total reading times for the five texts over the course of the experiment. Differences in reading times were investigated using a Repeated Measures MANOVA. The results show that there were significant differences in reading times for the texts, $F(4, 30) = 17.11$, $p < .05$. Pairwise comparisons (with Bonferroni corrections) showed that reading times for Block 1 differed from all other blocks, which were not different from each other. Thus, although there does appear to be an increase in reading times once participants were instructed to use a specific reading strategy, there did not appear to be any further systematic increase in reading time over the remainder of the experiment. Figure 3 shows the means of these reading times. Appendix J shows the reading times for each text by participant.

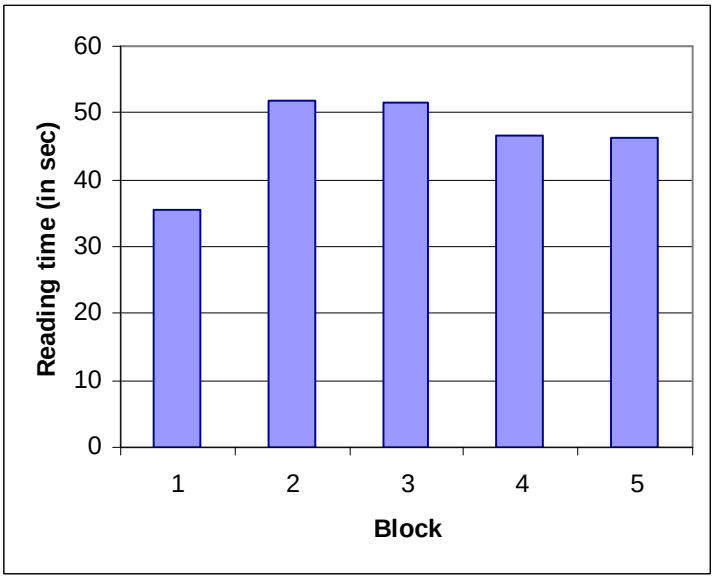

Figure 3.
Reading Times Over Eye Tracking Blocks

Block 1 was the Baseline task, Block 2 was the Main Idea Strategy Use task, Block 3 was the Anaphor Strategy Use task, Block 4 was the 1st Strategy Choice task that pulled for use of the main idea strategy, and Block 5 was the 2nd strategy choice task that pulled for use of the anaphor strategy.

The second validity analysis examined gaze time for eight randomly selected nouns (control words) in each text. This analysis was completed in order to determine whether the gaze time for the words associated with strategy instruction was longer than for randomly selected words from the text. In other words, this analysis attempted to determine whether longer gaze times were specific to the strategy used. Gaze time for these eight randomly-selected nouns for each block was correlated with strategy-consistent gaze times for each block. Across the blocks (1, 2, 3, 4 & 5) the only significant correlation between gaze time for the randomly-selected words and for words consistent with strategy use was for Block four, with a Pearson correlation of -.42 (N=34, $p<.05$). Thus, it does not appear that gaze time for strategy-consistent words is

associated to increased gaze time for other, randomly-selected words in the texts. Figure 4 shows the mean values for the strategy consistent, strategy inconsistent, and control words over the eye tracking tasks. As can be seen from the graph, gaze time over blocks is quite variable.

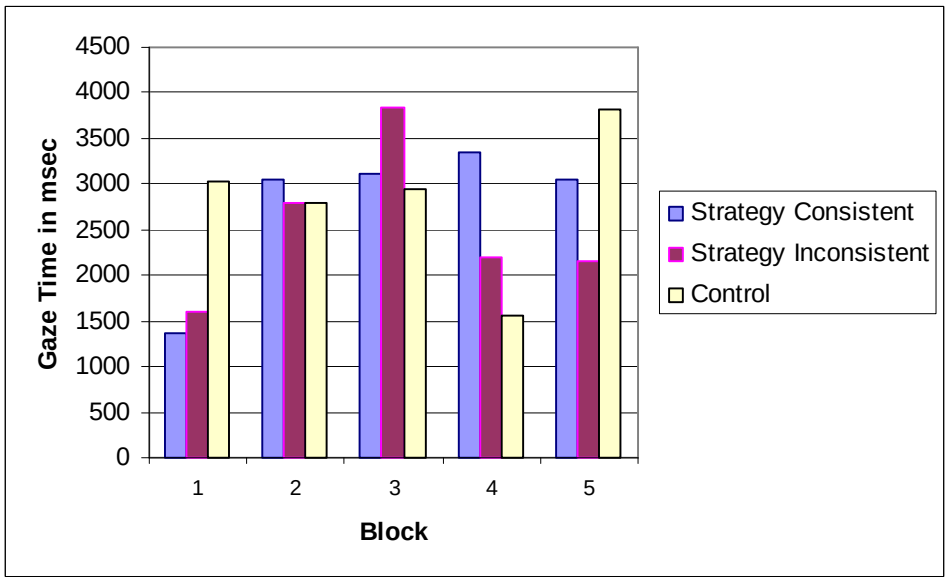

Figure 4.
Strategy Consistent, Strategy Inconsistent, and Control Word Gaze Times for the Eye Tracking Blocks

Block 1 was the Baseline task, Block 2 was the Main Idea Strategy Use task, Block 3 was the Anaphor Strategy Use task, Block 4 was the 1st Strategy Choice task that pulled for use of the main idea strategy, and Block 5 was the 2nd strategy choice task that pulled for use of the anaphor strategy.

The third validity analysis examined whether gaze time to main idea and anaphor words in the texts was associated with the strategy participants were asked to use, or was more indicative of general arousal. This was done by examining gaze time to words consistent with strategy instruction (i.e. gaze time to main idea words when the participant had received a main idea strategy instruction) and gaze time to words inconsistent with strategy instruction (i.e. gaze

time to anaphor words when the participant had received a main idea strategy instruction). It was expected that gaze time to words consistent with the strategy instruction would be higher than gaze time to words inconsistent with strategy instruction if the gaze time was associated with the use of a particular strategy, and not with general arousal. In order to investigate this question, ratios were constructed of gaze time to words consistent vs. inconsistent with the strategy that participants were asked to use. These ratios divided the gaze time to words consistent with the strategy participants were asked to use by the gaze time to words inconsistent with the strategy participants were asked to use. A higher ratio indicates a higher gaze time to strategy-consistent words. The means of these ratios are shown in Table 10.

Table 10.
Ratios of Gaze Time to Words Consistent vs. Inconsistent with Strategy Instruction

Block	Consistent Strategy Mean	Inconsistent Strategy Mean	Ratio
2	3040	2801	1.08
3	3121	3836	0.81
4	3022	1618	1.87
5	3374	2722	1.24

Table 10 shows that the gaze time for words consistent with the strategy is higher than the gaze time for words inconsistent with the strategy for all blocks except for Block 3 (anaphor strategy instruction pair). Table 10 also shows that the ratios of consistent to inconsistent strategy use are higher for the trials in which participants were asked to *choose* a strategy that was most useful for the particular text (main idea strategy for Block 4 and anaphor strategy for Block 5) than when they were instructed to *use* a particular strategy (main idea for Block 2 and anaphor for Block 3). These differences were investigated for statistical significance in Hypothesis 3.

The fourth validity analysis examined whether there was evidence that set-switching had occurred between the eye tracking texts. A Repeated Measures MANOVA was used to investigate this question. A total of ten variables were entered into the equation: gaze times for each for the five block conditions for both anaphor key words and main idea key words. For this analysis, a set switch was defined as a significant difference in gaze time to either main idea or anaphor words between two consecutive texts. The results show that there were significant differences between the gaze times for different blocks [$F(9,24) = 13.73$, $p<.05$]. Ten planned comparisons (with Bonferroni corrections) were conducted to determine whether there were significant differences between block pairs. Table 11 shows data for gaze time to main idea strategy words for these block pairs. Table 12 shows data for gaze time to anaphor strategy words for these block pairs.

Table 11.
Main Idea Strategy Word Gaze Times for Block Pairs

Blocks	Mean Diff (1^{st}-2^{nd})	*p*
1 to 2	-1804	<.05*
1 to 3	-2600	<.05*
2 to 3	-795	.99
3 to 4	814	.99
4 to 5	300	.99

* significant at p<.05

Table 12.
Anaphor Strategy Word Gaze Times for Trial Pairs

Trials	Mean Diff (1^{st}-2^{nd})	*p*
1 to 2	-1071	.05
1 to 3	-1392	<.05*
2 to 3	-320	.99
3 to 4	1502	<.05*
4 to 5	-1755	<.05*

* significant at p<.05

Tables 11 and 12 show that there were significant set switches on 5/10 possible switch pairs. The expected pattern of increased gaze times to strategy words after baseline was supported by these data. However, this increase in gaze time was not specific to the strategy being targeted by the task; increases occurred for main idea as well as anaphor words for both Blocks 2 and 3. Second, the expected patterns of significantly decreased gaze times for the main idea strategy words from Block 2 (main idea strategy) to Block 3 (anaphor strategy) and significantly increased gaze times for the anaphor strategy words from Block 2 to Block 3 were not supported by the data.

Most consistent with expectations was that gaze time for anaphor words decreased significantly between Blocks 2 (anaphor strategy) and 4 (main idea strategy) and increased significantly between Blocks 4 (main idea strategy) and 5 (anaphor strategy). Findings inconsistent with expectations were that gaze time for main idea words did not change significantly either between Blocks 3 and 4 or between Blocks 4 and 5.

Research Hypotheses

In the analysis of the final study results, all subtests of the standardized neuropsychological measures were analyzed separately, instead of examining summed scores, unless indicated. A number of correlations were calculated in order to determine whether the hypotheses were supported.

Hypothesis 1: It was predicted that scores on set maintenance as assessed by executive function measures would be significantly negatively correlated with the number of training trials necessary to reach learning criterion for each new strategy taught.

To test this hypothesis, the neuropsychological test measures of set maintenance (TMT Conditions 2 and 3, CWIT Conditions 1 and 2, BCT Maintenance, and CPT Omissions) were correlated with the score on the training tasks for each new strategy learned. A higher score on the training tasks indicated better performance on these tasks. The maintenance score on the BCT was the maximum number of items correct in a row, with a higher score indicating better maintenance skills.

Tables 13 and 14 show the results of these correlations. The main idea strategy training score was significantly positively correlated [$r(34) = .40$, $p < .05$] with Trail Making Test Condition 3 scaled score (letter sequencing). All other correlations with the other measures on

the Trail-Making Test, Color-Word Test, and the Booklet Category Test were not significant. Thus, this hypothesis was not well-supported.

Table 13.
Correlations of Executive Function Maintenance Measures with Main Idea Training Task Score

	r	*p*	n
TMT 2	.29	.09	34
TMT 3	.40	<.05*	34
CWIT 1	.04	.82	34
CWIT 2	.23	.18	34
BCT Main	.12	.50	34
CPT Omit	.13	.47	34

* $p < .05$

Table 14.
Correlations of Executive Function Maintenance Measures with Anaphor Training Task Score

	r	*p*	n
TMT 2	.08	.67	34
TMT 3	.05	.80	34
CWIT 1	-.07	.69	34
CWIT 2	.24	.18	34
BCT Main	.18	.30	34
CPT Omit	.24	.17	34

Hypothesis 2: It was predicted that set maintenance as assessed by executive function tests, and vigilance, as assessed by the attention task, would be significantly and positively correlated with set maintenance as assessed within the reading comprehension paradigm. Similarly, set switching as assessed by executive function tests would be significantly and positively correlated with set switching as assessed within the reading comprehension paradigm.

In order to test the first part of this hypothesis (set maintenance), maintenance measures on the executive function measures were correlated with the maintenance score on the strategy tasks (maintenance ratio for anaphors and maintenance ratio for main ideas). TMT Conditions 2 and 3, CWIT conditions 1 and 2, BCT Maintenance, and CPT % Omissions scores were used to assess the executive functions of maintenance. The maintenance score on the BCT was the number of items correct in a row, with a higher score indicating better maintenance skills. The maintenance score of the eye tracking texts was the absolute value of the change in gaze time between two trials in which the same strategy was to be used. A lower absolute difference score was expected to indicate better maintenance of the strategy over time, with the second score expected to be similar to the first score.

No significant correlations were found between the executive function measures of set maintenance and the anaphor and main idea set maintenance ratios from the strategy use tasks. Thus, the hypothesis was not well-supported. Tables 15 and 16 show these results.

Table 15.
Correlations of Executive Function Measures with Eye Tracking Maintenance Score for Texts 2-3

	r	*p*	n
TMT 2	.28	.11	34
TMT 3	.25	.15	34
CWIT 1	-.24	.17	34
CWIT 2	.01	.99	34
BCT Main	-.11	.53	34
CPT Omit	-.02	.93	34

Table 16.
Correlations of Executive Function Measures with Eye Tracking Maintenance Score for Texts 4-5

	r	*p*	n
TMT 2	-.07	.71	34
TMT 3	-.07	.70	34
CWIT 1	-.10	.56	34
CWIT 2	-.22	.22	34
BCT Main	-.28	.11	34
CPT Omit	.16	.36	34

* $p < .05$

In order to test the second part of the hypothesis (set switching), set switching scores on the executive function measures were correlated with the set switching scores on the strategy

tasks. TMT condition 4 and CWIT condition 4 were used as measures of the executive function of switching. The switching score on the BCT was a dichotomized yes/no value as to whether a criterion of 10 in a row correct had been met for each subtest. This was chosen as a criteria for measuring switching because a string of correct responses would indicate that the participant effectively switched strategies from the previous subtest. The scores on the eye tracking tasks were dichotomized values of 1 or 0, with 1 indicating a correct switch between using the correct strategy in two consecutive trials, and 0 indicating that the participant did not use the correct strategy in one or both of the consecutive trials. These values of 1 and 0 were calculated from the switch ratios (consistent/inconsistent) that were previously computed.

Previous analyses showed that gaze time when participants were instructed to use the main idea and anaphor strategies (Blocks 2 and 3) was associated with more non-specific activation, in which all strategies were used at relatively equivalent levels, but that gaze time for Blocks 4 and 5, in which participants were asked to choose a strategy, was associated with greater strategy-consistent gaze times. Thus, set switching was measured using the values for Blocks 4 and 5.

A Repeated Measures MANOVA was conducted to determine if there were differences in executive function switching measures (TMT condition 4 and CWIT condition 4) between those who had a correct switch between Blocks 4 and 5 on the eye tracking tasks and those who did not. The results show that there were not significant differences in executive function measures for those who had a correct switch between Blocks 4 and 5 on the eye tracking texts and those who did not, $F(2,32) = .46$, $p = .50$.

The switching score on the BCT was dichotomized to yes/no values based on whether a criterion of 10 in a row correct had been met for each subtest. Thus, a Chi-Square test was used to determine if the BCT score differed significantly for the groups who did and did not switch correctly on the eye tracking texts. There was no significant difference between groups ($p = .08$).

Next, the planned covariates (TMT Condition 1, TMT Condition 5, CWIT Condition 3, Digit Span, Sentence Memory, and Passage Comprehension) were entered, one at a time, to calculate partial correlations between the executive function measures and the eye tracking strategy use measures. This set of analyses revealed that no correlations between the executive function measures and eye tracking measures remained significant, except when controlling for Digit Span. When controlling for this condition, the correlations between TMT Condition 2 and the eye tracking maintenance score between the original texts 2 and 3 [$r(34)=-.41$, $p<.05$)], and the correlation between TMT Condition 3 and the eye tracking maintenance score between the original texts 2 and 3 [$r(34)=.36$, $p<.05$)] remained significant.

Overall, there was limited support for this hypothesis.

Hypothesis 3: It was predicted that participants would respond to experimenter instructions and use the main idea strategy more than the anaphor strategy in the second block while using the anaphor strategy more than the main idea strategy in the third block of the experimental paradigm. They were also expected to rate these specific strategies as more used on the strategy survey. It was further predicted that participants would use specific strategies more consistent

with qualities of the text rather than strategies inconsistent with the qualities of the text in the fourth and fifth blocks of the experiment. Participants were expected to use the main idea strategy more and rate it as more useful for the fourth block than for the fifth block and use the anaphor strategy more and rate it as more useful for the fifth block than for the fourth block.

This hypothesis was tested by comparing the means of gaze times between the eye tracking texts. These means were calculated as ratios of gaze time to words consistent with the strategy over gaze time to words inconsistent with the strategy. If the hypothesis was supported, the ratio of consistent to inconsistent words would be higher for the main idea strategy for Block 2 and Block 4 and higher for the anaphor strategy for Block 3 and Block 5. Also, the participants' ratings of how much they used strategies and how useful they found them were compared for all blocks to determine whether the participants rated the strategies differently as to their perceived usefulness and amount of use.

First, gaze time to the main idea and anaphor strategy words was examined over the course of the entire experiment. A Repeated Measures MANOVA was used to determine whether the gaze times to main idea and anaphor words were significantly different within a block. Overall, there were differences in gaze times to main idea and anaphor words *between all blocks* $F(2,32) = 19.79$, $p < .05$. However, posthoc pairwise comparisons with Bonferroni corrections revealed that main idea and anaphor key word gaze times did not differ significantly *within blocks*: Block 1 (mean diff = -493.63, $p = .99$), Block 2 (mean diff = 239.45, $p = .99$), Block 3 (mean diff = -714.89, $p = .99$), Block 4 (mean diff = 1403.29, $p = .35$), or Block 5 (mean

diff = -632.28, p = .99). Figure 5 shows the gaze time for main idea and anaphor strategy key words over the course of the experiment.

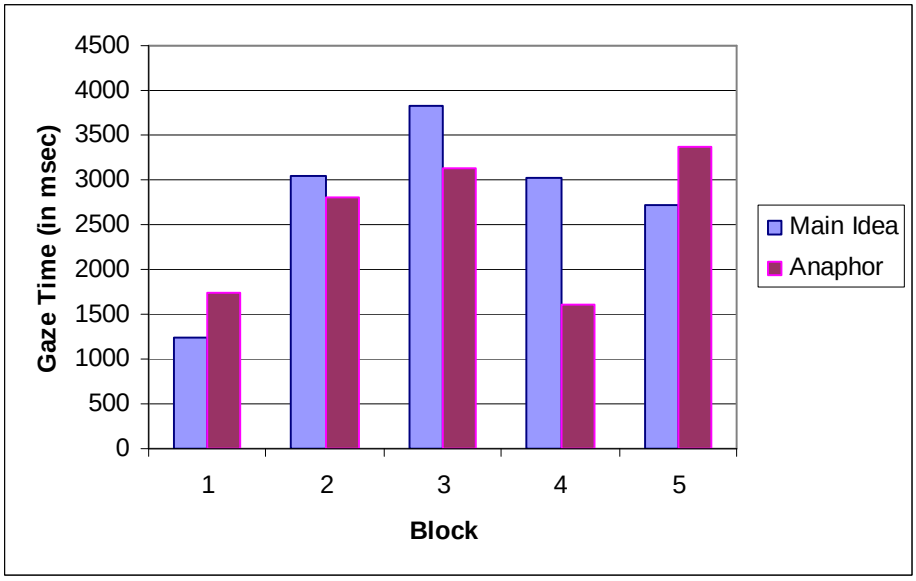

Figure 5.
Strategy Use Over Blocks for the Eye Tracking Task

Block 1 was the Baseline task, Block 2 was the Main Idea Strategy Use task, Block 3 was the Anaphor Strategy Use task, Block 4 was the 1st Strategy Choice task that pulled for use of the main idea strategy, and Block 5 was the 2nd strategy choice task that pulled for use of the anaphor strategy.

Thus, gaze times did not differ significantly between main idea words and anaphor words for any of the blocks. However, the graph shows an interesting dissociation of gaze patterns over the course of the experiment. Gaze time to both main idea and anaphor words increased from Block 1 for both anaphor and main idea strategies. However, a different pattern appeared for Blocks 4 and 5. Gaze time was higher for main idea words than for anaphor words for Block 4. Gaze time to anaphors was higher than gaze time to main idea words for Block 5.

These results, taken with the evidence of strategy-consistent use in Table 10, indicates that participants were likely to use appropriate strategies in a situation in which they were given a choice of strategies and in which characteristics of the text pulled for the use of a particular strategy. In contrast, gaze times were less consistent with strategy instructions when participants had just learned the strategy and were asked explicitly to apply it to a text that did not pull for use of a particular strategy.

The second part of hypothesis three (strategy survey) was examined by comparing the strategy survey responses using a Repeated Measures MANOVA. Planned comparisons were used to determine whether there were significant differences between participant ratings of strategy use and/or usefulness for the anaphor and main idea key words within blocks. Fourteen pairs of items were examined in the planned comparisons analysis, two for each original text (1) how much the anaphor vs. main idea strategy was used and (2) how useful the anaphor vs. main idea strategy was for the particular text. Only 4/14 comparisons showed significant differences between the strategy ratings within blocks. These significant findings were between the utility of strategies for Block 1 (anaphor strategy was rated as more useful), the use and utility of strategies for the first text in Block 2 (the anaphor strategy was rated as more used and more useful), and the utility of strategies for the second text in Block 2 (the anaphor strategy was more useful). Thus, participant ratings of strategy use did not necessarily match the participants' behavior.

Hypothesis 4: It was predicted that scores on set switching and maintenance on neuropsychological tests and from the eye tracking data would be significantly positively

correlated with participants' scores on a measure of self-reported use of reading comprehension strategies (Metacognitive Reading Strategies Questionnaire (MRSQ; Taraban, Kerr, & Rynearson, 2004).

To test this hypothesis, the scores from the MRSQ analytic and pragmatic strategy indexes were correlated with measures of set switching and maintenance from the neuropsychological tests. The same tests used to assess executive function set switching and maintenance in previous hypotheses were used to examine this hypothesis.

The MRSQ analytic strategy score was not significantly negatively correlated with any of the measures used to assess executive function set switching and maintenance. The MRSQ pragmatic strategy score was significantly negatively correlated with Omissions % T score [$r(34) = -.35$, $p < .05$]. None of the other measures used to assess executive function set switching and maintenance were significantly correlated with the MRSQ pragmatic strategy score. Thus, there was limited support for this hypothesis. Tables 17 and 18 show the results of these correlations.

Table 17.
Correlations of the MRSQ Analytic Strategy Score with Executive Function Measures

	r	*p*	n
TMT 2	-.06	.74	34
TMT 3	-.14	.43	34
TMT 4	.06	.73	34
CWIT 1	.27	.12	34
CWIT 2	.24	.17	34
CWIT 4	.19	.27	34
BCT Main	-.01	.97	34
CPT Omit	-.08	.66	34

Table 18.
Correlations of the MRSQ Pragmatic Strategy Score with Executive Function Measures

	r	*p*	n
TMT 2	-.04	.81	34
TMT 3	-.18	.32	34
TMT 4	.01	.98	34
CWIT 1	-.01	.95	34
CWIT 2	.06	.75	34
CWIT 4	-.10	.59	34
BCT Main	-.09	.61	34
CPT Omit	-.35	<.05*	34

* $p<.05$

Hypothesis 5: It was predicted that strategy use, as assessed by eye tracking data, would be significantly positively correlated with the participants' subjective judgments of when they used each strategy at each stage of the strategy use tasks (Strategy Use Survey).

In order to test this hypothesis, the scores obtained using the eye tracker for the amount of time main idea or anaphor words were viewed for each trial was correlated with the participant's rating, on a Likert scale of 1-5, of how much they used each strategy during each trial and how useful they found each strategy. Thus, there were sixteen correlations calculated to test this hypothesis - for main idea and anaphor use for text one (4 correlations total), and for the relevant strategy for the rest of the six texts (12 correlations total). A positive correlation between the more objective eye tracker data and the more subjective rating data was predicted.

There were 2/16 possible significant correlations between gaze time for the eye tracking tasks and self-report of strategy use on the Strategy Use Survey: for how much the anaphor strategy was used in the first anaphor strategy text [$r(34) = .52$, $p < .05$] and how much the anaphor strategy was used in the anaphor switch text [$r(34) = .38$, $p < .05$]. There were no other significant correlations. Thus, there was limited support for this hypothesis.

Chapter 4

Discussion

In this study, we examined the relationship between neuropsychological assessment and more real-world functioning by examining the relationship between executive functions, a cognitive domain that is commonly assessed in neuropsychological evaluations, and the use of metacognitive strategies, a common method of cognitive intervention. The current study attempted to determine the extent to which the executive functions (EF) of set maintenance and switching, as assessed by neuropsychological testing, can be used as predictors for the application of set maintenance and set switching during a more ecologically valid task. Specifically, this project examined how the executive functions of set maintenance and switching correlate with the ability to perform set maintenance and set switching while using two different metacognitive strategies during reading comprehension tasks. Set maintenance and switching, as assessed by executive function measures, was expected to be positively correlated with, and significantly predict, the ability of individuals to complete set maintenance and set switching during the metacognitive strategy use task.

In the study, participants completed a demographic questionnaire, and then received training in two different metacognitive reading comprehension strategies and read several texts silently while their eye movements were being monitored by an eye tracking device. Strategy use on the reading tasks was assessed by calculating the total gaze time on key words read during the eye tracking task, with the underlying assumption that longer gaze times on key words was a proxy indice of a particular strategy use. Measures of reading comprehension for the texts, a self-

report of strategy use, and a validity check on strategy use were completed after the eye tracking portion of the experiment was complete. Next, a vision screening and standardized measures of executive functions, attention, working memory, reading comprehension, and metacognitive strategy use during typical college reading were administered.

There were several blocks in the eye tracking portion of the experiment. Gaze times to main idea (i.e. words related to the main idea of the text) and anaphor (i.e. words that clarified pronouns in the text) key words were summed and used as measures of strategy use. Participants first completed a baseline task (Block 1) before they were taught any strategies. Then they were taught the main idea strategy and were asked to apply it to two texts (Block 2), then were taught the anaphor strategy and were asked to apply it to two texts (Block 3). In Blocks 4 and 5, participants were asked to choose the strategy that they found most useful in reading the texts. The texts for Blocks 4 and 5 were constructed so that one strategy would be more useful than another for reading and comprehending the text. There were two counterbalanced block orders. Participants were expected to maintain strategy use between texts when asked to apply the same strategy for consecutive trials. Participants were expected to switch strategy use in response to experimenter instructions or the demands of the texts.

A series of validity analyses was conducted to determine if the finding of increased strategy use with increased choice could be accounted for by other factors, such as general increased strategy use over time or increasing reading times over the course of the experiment. A Repeated Measures MANOVA was used to assess order effects for the eye tracking data. The two orders were not significantly different from each other. Thus, the data from the two

counterbalance orders were combined. Although there did appear to be an increase in overall reading times once participants were instructed to use a specific reading strategy, there did not appear to be any further systematic increase in reading time over the remainder of the experiment. Thus, participants did not simply take longer to read passages over the course of the experiment, resulting in longer gaze times, which could be mistaken for increased strategy use in the last few texts in the experiment. Also, analyses revealed that longer gaze times were specific to words associated with the strategy used and not to other, randomly-selected words in the texts. A third validity analysis determined that gaze time to main idea and anaphor words in the texts was associated with the strategy participants were asked to use, and could not be attributed to general arousal. Results revealed that the gaze time for words consistent with the strategy was higher than the gaze time for words inconsistent with the strategy for all blocks except for Block B (anaphor strategy instruction pair). Finally, set switching was shown occur for 5/10 possible pairs of texts. For these five texts, there were significant differences in gaze times between the two texts in a pair. For the other five texts, this statistically significant difference was not present. Thus, participants did not simply use the same strategy thoughout the experiment.

A number of hypotheses were proposed in this study. Set maintenance as assessed by executive function measures was expected to be significantly negatively correlated with the number of training tasks necessary to reach learning criterion for each new strategy learned. Also, set maintenance and set switching as assessed by executive function measures were expected to be positively correlated with set maintenance and set switching within the metacognitive strategy task. Also, it was predicted that participants would respond to experimenter instructions to use specific strategies more in the metacognitive strategy tasks.

They were also expected to rate these strategies as more used on the strategy survey. It was further predicted that participants would use strategies consistent with qualities of the text rather than strategies inconsistent with the qualities of the text during the metacognitive strategy portion of the experiment. Set switching and maintenance as assessed by neuropsychological tests and on the eye tracking tasks was expected to significantly correlate with participants' scores on a measure of self-reported use of reading comprehension strategies. Finally, strategy use as assessed by objective eye tracking measures was expected to be significantly correlated with the participants' subjective judgments of when they used each strategy at each stage of the strategy use tasks (Strategy Use Survey). None of these hypotheses was supported.

However, results did show an unexpected and very interesting relationship between experimenter directions and metacognitive strategy use. Participants were expected to use a particular strategy more when the experimenter asked them to utilize it for specific reading passages, as was done in Blocks 2 and 3. This hypothesis was not supported, with participants having similar gaze times to main idea and anaphor strategy key words in Blocks 2 and 3. However, participants did use the main idea strategy significantly more and rated it as more useful for Block 4 (which pulled for use of this strategy) than for Block 5 (which did not pull for its use) and used the anaphor strategy more and rate it as more useful for Block 5 (which pulled for use of this strategy) than Block 4 (which did not pull for its use). Although there were no significant differences between gaze time to main idea and anaphor words in Blocks 4 and 5, the ratios of consistent to inconsistent strategy use were higher for the trials in which participants were asked to choose a strategy that was most useful for the particular text (Blocks 4 and 5) than when they were instructed to use a particular strategy (Blocks 2 and 3). These results indicate

that participants were more likely to use appropriate strategies in a situation in which they were given a free choice of strategies and in which characteristics of the text pulled for the use of any particular strategy. In contrast, gaze times were less consistent with strategy instructions when participants had just learned the strategy and were asked explicitly to apply it to a text that did not pull for use of a particular strategy.

Overall, results suggested that there was not a clear relationship between normative test-based measures of set maintenance and set switching and participants' set maintenance and set switching within the context of a reading comprehension task. This study also suggests that participants' actual performance on tasks was not well-correlated with their self-report of their behaviors. However, participants did show increased use of the strategies over baseline levels, indicating that they did learn the strategies and applied them over the course of the study.

There is ample evidence that neuropsychological testing is a good predictor of everyday functioning in the areas of dementia (Bondi, Salmon, & Kaszniak, 1996), substance abuse (Teichner et al., 2001), spinal cord injury and TBI (Hanks et al., 1999; Ross et al., 1997; Crepeau & Scherzer, 1993; Bowman, 1996), HIV (van Gorp, Baerwalk, Ferrando et al., 1999), driving (Meyers, Volbrecht, & Kaster-Bundgaard, 1999) and return to work (Teasdale et al., 1997). However, a review of studies of the ecological validity of neuropsychological tests by Chaytor and Schmitter-Edgecombe (2003) found that the magnitude of the relationship between the tests and measures of everyday functioning was in the moderate range and that many individual tests were only weakly correlated to measures of outcome. Specifically, within the area of executive functions, they found that the standardized tests did not correlate well with measures of self-

report of everyday skills, but did correlate significantly with everyday skills as assessed by informants and clinician ratings. These results make sense when one considers that a common consequence of executive function deficits is decreased self-monitoring skills. Chaytor et al.'s (2003) finding of nonsignificant relationships between executive function tests and self-report measures is consistent with the findings of the current study that participants' ratings of their strategy use did not correlate well with their actual strategy use, at least as measured by eye tracking.

Chaytor and Schmitter-Edgecombe (2003) also found that relationships between neuropsychological tests and outcome measures were stronger when the outcome measure was more directly related to the domain measured by the executive functioning test. In the current study, the executive function measures of set switching and set maintenance were related to the demands of the metacognitive task (maintaining and switching set between texts), but did not specifically measure these skills within the areas of strategy learning or reading comprehension. The lack of a one-to-one correspondence between neuropsychological tests and measures of metacognitive strategy use was an expected part of this study (as executive function tests do not assess metacognition or reading comprehension specifically), but may have led to lower correlations that expected between these two domains. It may be that the ability to perform set maintenance and set switching executive function tasks is not necessarily correlated with one's actual performance of these tasks. Indeed, the executive function tasks used in this study were relatively well-defined and the stimuli used were not overly complex. On the other hand, the metacognitive strategy task involved elements of reading comprehension, strategy choice, and self-monitoring which added to the complexity of the task. It may be that the currently-available

tests of executive functioning cannot adequately capture the complexity of real-world situations which require the on-line application of executive function skills.

In this study, we did find that there was improved metacognitive strategy use in situations in which there was increased free choice of strategy/lack of specific instructions *and* elements within the text that guided the learner to use a particular strategy. Future studies in this area should attempt to determine whether it is increased choice/lack of specific instructions *or* elements of structure attributes within the texts that best account for these findings, or an interaction of the two.

Increased choice may have a positive effect because it may increase the learner's motivation to perform the task. Indeed, personal motivational states have been theorized to help determine the course of strategy acquisition, the likelihood of strategy transfer, and the quality of understanding about the role of mental processes (Borkowski, Carr, Rellinger, & Pressley, 1990). Also, as metacognitive strategies are thought to require conscious control, it is not surprising that increased engagement in a task might elicit more appropriate use of these strategies.

Elements of text attribute structure may have also accounted for increased correct strategy use for texts in which participants were given a choice of strategies. In order to encourage participants to use a strategy, the texts were made more difficult in a particular way to make a particular strategy more useful in understanding the text. For the text which "pulled" participants to use the main idea strategy, the main idea was "buried" in the middle of the text and not explicitly stated. For the text that "pulled" for the anaphor strategy, there were several

inconsistent anaphors (i.e. a female pronoun would be replaced with a male pronoun). These changes to the texts were expected to make them more salient to readers and encourage the use of a particular strategy, which is exactly the result that was seen in this study. Similar findings of increased task performance with increased stimulus saliency have been observed in individuals from preschool through college age. For example, in one study (Vlietstra, 1978), preschoolers were trained in a selective attending strategy to match pictures from memory. For one third of the participants, relevant portions of the stimuli were made perceptually salient, for another third, irrelevant portion were made salient, and for the last third, no portions of the stimuli were made salient. Strategy training was found to increase performance on the task, and irrelevant perceptual cues were found to interfere with performance of the task. A developmental study by Eimas (1970) examined matrix problem-solving in elementary and middle school-aged children and college students. Better performance on the matrix task was associated with increased saliency of the matrix designs (defined as number of categorical responses that could describe a stimulus) as well as increasing age of the participants.

Similarly, stimulus equivalence has been found to be very important in teaching new information. Stimulus equivalence occurs when two or more stimuli possess symmetry (one is selected as readily as the other with both stimuli present), transitivity (selecting a new stimulus with an already learned stimulus present without direct training), and reflexivity (selecting an identical stimulus to one already learned without direct training) (de Rose, de Souza, & Hanna, 1996). Equivalence classes can include all stimuli that have been shown to be equivalent through testing, including the picture of a word, the written word, and the dictated word. De Rose al al. (1996) taught seven first-grade students who did not yet read how to match printed and dictated

words and how to copy printed words with movable letters and name them. The children learned 51 new words using this paradigm. However, they were not consistently able to generalize their learning to new, untaught words. In a follow up study, Melchiori, de Souza, and de Rose (2000) found that by recombining the syllables of the training words, participants were able to generalize the results of training trials to new words. These results suggest that there must be a very close relationship between two stimuli in order for learning to transfer. Individuals in the current study were able to transfer the information acquired in the strategy training portion of the experiment to the strategy usage (eye tracking) portion of the experiment. However, this study did not include individuals with very poor executive function as might be seen in an acute medical population. It may be that such individuals with poor executive function skills may require a more structured approach to learning that the one provided in this experiment, including equivalence-based methods that change the stimuli in a structured, step-by-step manner.

In the current study, we looked at the relationship between two specific aspects of executive functions and metacognition, set switching and set maintenance. There is evidence that more general measures of executive functions correlate well with behavior for those with neurological impairment. However, the current study did not examine individuals with neurological impairment. It is certainly conceivable that the relationship between executive functions and metacognitive skills may be different for a neurologically typical population. Another reason that this study did not find a significant relationship between executive functions and metacognition is that perhaps executive function needs to be examined as a unified construct, not as a series of subskills which make up executive function. Set switching and set

maintenance may not be completely dissociable concepts from the rest of executive function. Finally, the task demands for the executive function tasks and the metacognitive tasks were different. The executive function tasks [DKEFS CWIT and Trails (Delis, Kaplan, & Kramer, 2001), CPT (Connors, 1992), BCT (DeFilippis & McCampbell, 1997)] measure a narrow set of skills (i.e. connecting dots in order for the trails), but use this narrow sample of behavior to generalize to a much wider construct (i.e. set maintenance). In contrast, for the metacognitive strategy task, the stated task demand was exactly the same as what was measured (i.e. looking at the key words). It is likely that a task with a one-to-one correspondance between task demand and measured skill (metacognitive strategy use) and a task which has a generalized dependent variable (EF) may not be highly correlated. In addition, the metacognitive strategy task was infused with context (i.e. the participants were reading stories) and the goal of the task was clear (look at the key words, understand the passage). The task was familiar (read the passage) and the strategies that needed to be used were trained and clearly articulated. Contrast this with the executive function tasks, in which participants were presented with unfamiliar tasks and given sometimes very vague instructions about how to complete the task (i.e. Booklet Category task). The mismatch between the structure and demands of the executive function vs. metacognitive tasks could help explain why no significant relationship was found between the two.

The results of this study suggested that standardized measures of set maintenance and set switching did not predict participants' abilities to perform set maintenance and set switching while they used different metacognitive strategies. Perhaps standardized neuropsychological testing is not the best way to predict whether an individual will have success at learning or applying metacognitive strategies in applied situations. However, this study did suggest that

adding measures of possible components of the metacognitive strategy may increase the ability to predict metacognitive strategy use from executive function scores. When planned covariates [TMT Condition 1, TMT Condition 5, CWIT Condition 3 (Delis, Kaplan, & Kramer, 2001), Digit Span (Wechsler, 1997), Sentence Memory (Thorndike, Hagen, & Sattler, 1987), and Passage Comprehension (Woodcock, McGrew, & Mather, 2001)] were entered to calculate partial correlations between the executive function measures and the eye tracking strategy use measures no correlations between the executive function measures and eye tracking measures remained significant, except when controlling for CWIT Condition 3. These results suggest that adding measures of working memory, reading comprehension, visual scanning, and motor speed to neuropsychological assessments may help to capture some of these "real-world" factors associated with metacognitive strategy use.

Applying metacognitive strategies to a real text is a substantially different experience from completing a series of neuropsychological tests that assess the factors that are hypothesized to contribute to this task. The metacognitive strategy task required the participant to use a number of skills in parallel, including reading comprehension, awareness of one's level of comprehension of the task, remembering to use the strategies taught, actually using the strategies taught, familiarity with the topic of the text, reading speed, motivation to use the strategies and read critically, adjusting to reading using an eye tracking device, text variables such as reading level and number of nouns, and a number of other cognitive factors that are difficult to predict and measure. Additionally, choice of strategy and text structure attributes were shown to be important variables in this study, with participants using the correct strategy more often when they were given a choice of strategy and when the text was structured in a particular way to

"pull" for a certain strategy. The executive function tasks assess only a small portion of these behaviors. Given that any task in the natural environment will be necessarily different from performing a similar task in an experimental testing environment, it may be more appropriate to conceptualize the relationship between assessment and intervention as containing a number of uncontrolled variables. Perhaps a certain number of these variables can be accounted for, and the rest can be controlled, at least temporarily in the early stages of intervention, by providing environmental structure. For example, if an individual shows difficulties with reading comprehension and executive function, perhaps a metacognitive intervention could begin with a great deal of structure, such as having a notecard available reminding the individual to use the strategy, providing cues or structure in the text to use the strategy, prompting with a tone or other reminder to use the strategy while reading the text, using serial eye tracking to measure any increases in gaze time to key words, and other methods. Perhaps over time this external structure can be faded and the individual can internalize some of the prompts. The work on stimulus equivalency reviewed indicates that very small stimulus changes from trial to trial are often needed for learning, and especially for generalization to occur. As the individual generalizes the behavior, the structure will necessarily need to be removed to adapt to the variety of situations the individual may encounter.

A more adaptive approach to assessment may be appropriate when its purpose is to determine how an individual is likely to respond to a particular intervention. This could involve presenting the individual with a series of tasks with differing levels of environmental structure in the testing situation or "taking the assessment on the road" and observing the individual in the natural environment. In this way, assessment of everyday performance could be an important

first step in intervention. This approach differs from response to intervention as practiced in schools because it would be a brief addition to the assessment experience, not an instructional program delivered by the school. Using an adaptive assessment approach, individuals could be placed more immediately in the correct type of intervention, with the correct amount of environmental structure, based on their test results in adaptive assessment. For example, the Dynamic Assessment of Test Accomodations (DATA; Fuchs et al., 2003) assesses how individuals perform in testing situations with different test structures (timed, untimed, with breaks). This model could be applied to many different cognitive skills for more effective assessment and recommendations that specify how much environmental structure and types of supports that would be needed for an individual to achieve their optimal learning and everyday performance. By using adaptive assessment methods as an integral part of assessment, neuropsychology could help to close the gap between assessment and intervention and allow us to prove, not just guess, what effect a proposed intervention could have on performance.

References

Anderson, P. (2002). Assessment and development of executive function during childhood. *Child Neuropsychology, 8*(2), 71-82.

Barkley, R.A. (1996). Linkages between attention and executive functions. In G.R. Lyon & N.A. Krasneger (Eds.), *Attention, Memory, and Executive Function* (pp. 307-325). Baltimore, MD: Paul H. Brookes Publishing Co., Inc.

Bondi, M.W., Salmon, D.P., & Kaszniak, A.W. (1996). The neuropsychology of dementia. In I. Grant & K.M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders (2nd ed.)*. New York: Oxford University Press.

Borkowski, J.G., Carr, M., Rellinger, E., & Pressley, M. (1990). Self-regulated cognition: Interdependence of metacogntion, attributions, and self-esteem. In B.F. Jones & L. Idol (Eds.), *Promoting Academic Competence and Literacy in School* (pp. 477- 501). San Diego, CA: Academic.

Bowman, M.L. (1996). Ecological validity of neuropsychological and other predictors following head injury. *The Clinical Neuropsychologist, 10*, 382-396.

Burke, J.M., Smith, S.A., & Imhoff, C.L. (1989). The response styles of post-acutre brain-injured patients on MMPI. *Brain Injury, 3*, 35-40.

Carpenter, P.A., & Just, M.A. (1978). Eye fixations during mental rotation. In J.W. Senders, D.F. Fisher, & R. A. Monty (Eds.), *Eye movements and higher psychological functions*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Chaytor, N. & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review, 13*(4), 181-197.

Cohen, J. (1977). *Statistical power analysis for behavioral sciences (rev. ed.)*. New York: Academic Press (Reprinted edition 1988 Hillsdale, NJ: Erlbum).

Conners, C.K. (1992). *Conners' Continuous Performance Test*. Toronto: Multi-Health Systems.

Crepeau, F. & Scherzer, P. (1993). Predictors and indicators of work status after traumatic brain injury: A meta-analysis. *Neuropsychological Rehabilitation, 3*, 5-35.

DeFilippis, N.A. & McCampbell, E. (1997). *Booklet Category Test (BCT) (2nd ed.)* Odessa, FL: Psychological Assessment Resources.

Delis, D.C., Kaplan, E., & Kramer, J.H. (2001). *Delis-Kaplan Executive Function System (D-KEFS): Examiner's Manual*. San Antonio: The Psychological Corporation.

de Rose, J.C., de Souza, D.G., & Hanna, E.S. (1996). Teaching reading and spelling: Exclusion and stimulus equivalence. *Journal of Applied Behavior Analysis, 29*(4), 451-469.

Diamond, A., Prevor, M.B., Callender, G., & Druin, D.P. (1997). *Prefrontal cortex cognitive deficits in children treated early and continuously for PKU*. Monographs of the Society for Research in Child Development, 62(4), 1-205.

Dupont, W.D. & Plummer, W.D. (1997). *PS Power.* Controlled Clinical Trials, 18, 274.

Ehrlich, K. & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and the immediacy or processing. *Journal of Verbal Learning and Verbal Behavior, 22*, 75-87.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20*, 641-655.

Eimas, P.D. (1970). Information processing in problem solving as a function of developmental level and stimulus saliency. *Developmental Psychology, 2*(2), 224-229.

Flavell, J.H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906-911.

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

Frazier, L. & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14,* 178-210.

Fuchs, L., Fuchs, D, Eaton, S. & Hamlett, C. (2003). *Dynamic Assessment of Test Accomodations (DATA).* Harcourt Assessment, Inc: San Antonio, TX.

Hanks, R.A., Rapport, L.J., Millis, S.R., & Deshpande, S.A. (1999). Measures of executive functioning as predictors of functional ability and social integration in a rehabilitation sample. *Archives of Physical Medicine and Remediation, 80,* 1030-1037.

Heaton, R.K., Miller, S.W., Taylor, M.J., & Grant, I. (2004). *Revised Comprehensive Norms for an Expanded Halstead-Reitan Battery: Demographically Adjusted Neuropsychological Norms for African American and Caucasian Adults*. Lutz, FL: Psychological Assessment Resources, Inc.

Just, M.A. & Carpenter, P.A. (1984). Using eye fixations to study reading comprehension. In D. Kieras & M. Just (Eds.), *New Methods in Reading Comprehension Research* (pp.151-182). Hillsdale, NJ: Lawrence Erlbaum Associates.

Just, M.A., Carpenter, P.A., & Woolley, J.D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111*, 228-238.

Kucera-Thompson, J. (2003). Executive processes and performance on the Booklet Category Test. *Dissertation Abstracts International: Section B: The Sciences & Engineering, 63*(8-B), 3924.

Lezak, M.D., Howieson, D.B., & Loring, D.W. (2004). *Neuropsychological Assessment*. Oxford University Press, New York.

Lodico, M. G. (1983). The effects of strategy-monitoring training on children's selection of effective memory strategies. *Journal of Experimental Child Psychology, 35*(2), 263-277.

Melchiori, L.E., de Souza, D.G., & de Rose, J.C. (2000). Reading, equivalence, and recombination of units: A replication with students with different learning histories. *Journal of Applied Behavior Analysis, 33*(1), 97-100.

Meyers, J.E., Volbrecht, M., & Kaster-Bundgaard, J. (1999). Driving is more than pedal pushing. *Applied Neuropsychology, 6*, 154-164.

Mullen, J.L. (1987). *College Reading and Learning Skills*. Englewood Cliffs, NJ: Prentice-Hall.

Pressley, M. (2002). Comprehension Strategies Instruction: A Turn of the Century Report. In C.C. Block & M. Pressley (Eds.). *Comprehension Instruction: Research-Based Best Practices* (pp.11-27). New York, NY: Guilford Press.

Rayer, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41*, 211-236.

Rayner, K. & Bertera, J.H. (1979). Reading without a fovea. *Science, 206*, 468-469.

Rayer, K. & Carroll, P.J. (1984). Eye movements and reading comprehension. In D. Kieras & M. Just (Eds.), *New Methods in Reading Comprehension Research* (pp. 129-150). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rayner, K., Pollatsek, A., Starr, M.S. *(2003).* In A. F. Healy & R.W. Proctor (Eds.), *Handbook of psychology: Experimental psychology, 4* (pp. 549-574). Hoboken, NJ: John Wiley & Sons, Inc.

Rayner, K., Well, A.D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception & Psychophysics, 27*, 537-544.

Ross, S.R, Millis, S.R., & Rosenthal, M. (1997). Neuropsychological prediction of psychosocial outcome after traumatic brain injury. *Applied Neuropsychology, 4*, 165-170.

Sbordone, R. & Guilmetter, T. (1999). Ecological validity: Prediction of everyday and vocational functioning from neuropsychological test data. In J.J. Sweet (Ed.). *Forensic Neuropsychology: Fundamentals in Practice* (pp. 227-254). Lisse, Netherlands: Swets & Zeitlinger.

Schneider, W. (1985). Developmental trends in the metamemory-memory behavior relationship: An integrative review. In D.L. Forrest-Presley, G.E. MacKinnon, & T.G. Waller (Eds.), *Metacognition, Cognition, and Human Performance, 1*, pp. 57-109. New York: Academic.

Smith, B.D. (2005). *Bridging the Gap: College Reading*. Needham Heights, MA: Pearson.

Sohlberg, M.M. & Mateer, C.A. (1989). *Introduction to Cognitive Rehabilitation: Theory and Practice*. New York, NY: Guilford Press.

Teasdale, T.W., Skovdahl Hansen, H.S., Gade, A., & Christensen, A.L. (1997). Neuropsychological test scores before and after brain-injury rehabilitation in relation to return to employment. *Neuropsychological Rehabilitation, 7*, 23-42.

Teichner, G., Horner, M.D., & Harvey, R.T. (2001). Neuropsychological predictors of the attainment of treatment objectives in substance abuse patients. *International Journal of Neuroscience, 106*, 253-263.

Thorndike, R.L., Hagen, E.P., & Sattler, J.M. (1987). *Stanford-Binet Intelligence Scale (4th ed.).* Chicago: Riverside.

Trabasso, T. & Bouchard, E. (2002). Teaching readers how to comprehend text strategically. In C.C. Block & M. Pressley (Eds.). *Comprehension Instruction: Research-Based Best Practices* (pp.176-200). New York, NY: Guilford Press.

Van Gorp, W.G., Baerwald, J.P., Ferrando, S.J., et al. (1999). The relationship between employment and neuropsychological impairment in HIV infection. *Journal of the International Neuropsychological Society, 5*, 534-539.

Vlietstra, A.G. (1978). The effect of strategy training and stimulus saliency on attention and recognition in preschoolers. *Journal of Experimental Child Psychology, 25*(1), 17-32.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-III*. San Antonio: The Psychological Corporation.

Woodcock, R.W., McGrew, K.W., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside.

Zola, D. (1982). The perception of words in reading: An exploratory treatise. *Dissertation Abstracts International, 42*(12-B, Pt 1), 4959-4960.

Appendices

Appendix A

Training Instructions for Main Idea Strategy

Modified from Mullen (1987)

The ability to find the main idea is important for comprehension and also for note-taking from both lectures and textbooks. If you can isolate main ideas, you are able to focus on the overall theme and the most important points of what you are reading or hearing. In this section you will learn how to find the main idea of paragraphs.

The main idea may be found anywhere in a passage. It may even be unstated (implied) throughout the material. To find the main idea of a paragraph:

1. Ask yourself what the paragraph told you.
2. Think about what the paragraph told you and decide which points are narrow in scope. These are points, details, that relate to just one part of the paragraph.
3. Decide which points are more global or general in scope. These are points that the rest of the paragraph can relate to and are usually the basis for determining the main idea.
4. If there is just one general point, it is the main idea. You should be able to take this point out of the material as a complete thought or else reword it so that it is one.
5. If there is more than one general point, decide how they relate to each other. This stated relationship will be the main idea.

6. If there is no general point, think about what it might be. This would be the implied main idea.
7. Most often, but not always, the main idea of a textbook paragraph is found in the first or second sentence.

Position of Main Idea in a Paragraph

First sentence:

As I grow older, autumn becomes my favorite time of year. The reprieve from summer's heat is welcomed by both my mind and body – by a body for physical comfort and by a mind beginning to grow restless from summer's languor. The promise of another academic year adds to the optimism I feel at the start of the season.

Middle sentence:

Autumn's reprieve from the heat of summer is welcomed by both my mind and body – by a body for physical comfort and by a mind beginning to grow restless from summer's languor. As I've grown older, this reprieve has made autumn my favorite time of the year. This, combined with the promise of another academic year, causes the optimism I feel at the start of this season.

(Note: When the main idea is in the second sentence, the first sentence is often a transition or an introduction.)

Last sentence:

Autumn's reprieve from the heat of summer is welcomed by both my mind and body – by a body for physical comfort and by a mind beginning to grow restless from summer's languor. This relief, combined with the promise of another academic year, causes the optimism I feel. For these reasons, as I grow older, autumn has become my favorite season.

Split (first and last sentence):

As summer ends, my favorite season approaches. Autumn's reprieve from the heat of summer is welcomed by both my mind and body – by a body for physical comfort and by a mind beginning to grow restless from summer's languor. The promise of another academic year adds to my optimism. Indeed, as I grow older, autumn has become my favorite season.

Unstated:

What season do I favor? Autumn's reprieve from the heat of summer is welcomed by both my mind and body – by a body for physical comfort and by a mind beginning to grow restless from summer's languor. This and the promise of another academic year add to the optimism I feel at the start of this season.

You can test to see if your selected main idea is broad enough. If all parts of the paragraph relate to your statement, you have keyed in on the main idea. You may have to narrow it down. When it can no longer be narrowed and still include all the paragraph's parts, that is the main idea.

Examples

Read each of the following paragraphs and write the main idea in the space provided.

Paragraph 1:

In small, primitive societies, the family is largely self-sufficient, providing for most of its members' needs. The entire family cooperates to make tools, build shelters, and hunt, gather, or grow food. Parents and elders teach children the skills they will need as adults, as well as a sense of right and wrong. Religious practices are woven into the pattern of daily family lie. The head of the family decides who does what and settles disputes.

Main Idea: _____

Paragraph 2:

Long-term memory, then, is permanent. But how is it organized? There is some debate on this point, but most often long-term memory is compared to a library and its card catalog, or to a book and its index. Material is categorized or indexed. The card catalog or book index is then used to find what is needed. Similarly, information entering long-term memory is categorized or indexed according to its meaning. We can "look up" a piece of information by using the indexes. We may get to the word "Iowa" through thinking of corn, or hearing "Cedar Rapids" mentioned, or reading the word "Ionic" and recognizing the similarity of the initial sounds. The more indexes or associations an item has, the easier it will be to remember, just as it is easier to find a certain passage in a book if may of its key words and terms, rather than just one or two, are

indexed. This is one reason why we tend to remember semantic material better than episodic (Tulving, 1972).

Main Idea: _____

Paragraph 3:

It seems extremely important for young people to feel accepted by their peers. It also is important to them to be thought of as adult, or in control of their own lives. These two points, in combination, are reasons for teen-age drinking. Many teen-agers drink to take on the role of adults, and many others drink to be accepted by that group. Young people seem to regard drinking as a badge of adulthood, or virility. They also see it as a way to rebel against adults or society in general when they feel thwarted by those groups.

Main Idea: _____

Paragraph 4:

By the mid-twentieth century, it was obvious that many school systems were not doing their job – large percentages of the men drafted to fight in the two world wars were illiterate. At the same time, the nation's population was growing at a rapid rate as a result of the post-World War II baby boom, while teachers were moving to other fields because their salaries were too low. And to top it all off, in 1957 the Soviet Union launched Sputnik I, the first man-man satellite to orbit

the earth. Suddenly it became clear to large numbers of Americans that education and national security were related. The time for federal aid to education had come.

Main Idea: _____

If a perfect score is not obtained on the previous 4 texts, the next 4 texts will be given.

Additional Examples

Read each of the following paragraphs and write the main idea in the space provided.

Paragraph 1**:**

Leisure – meaning free or discretionary time – has expanded dramatically in recent decades. The causes are obvious: the shortened work-week, the increased number of holidays, longer vacations and longer periods of retirement through Social Security and other pension plans. The availability of such labor-saving devices as automatic washers and power movers has also simplified our lives by reducing the time needed to carry out many of the tasks of daily living.

Main Idea: _____

Paragraph 2:

Xenophobia, or antiforeign feeling, after the war brought to a head the anti-immigration sentiment that had been growing in the United States since the 1880's. The Immigration Restriction League had been organized by a group of New England intellectuals in 1894. It reflected fears that the "new" immigrants from southern and eastern Europe would destroy the "American character." This group had pushed for a literacy test for immigrants. During the Progressive era, xenophobes and labor and business leaders who shared hostility to newcomers were joined by liberals who feared that immigration was threatening the American way of life.

Main Idea: _____

Paragraph 3:

The third dimension of social stratification is prestige: the favorable evaluation and social recognition that a person receives from others. Prestige comes in many forms: public acceptance and fame, respect and admiration, honor and esteem. And it can the gain in many ways. People who are very kind, generous, brave, creative, or intelligent are often rewarded with prestige. Money can buy prestige, and power can demand it, or at least its outward appearance. For example, when John D. Rockefeller, Sr., made his first millions in oil, he was publicly despised. With the passage of time, however, he used his great wealth to gain prestige not only for himself but also for his heirs, but funding museums, parks, foundations, and charities.

Main Idea: _____

Paragraph 4:

Recreation, however, is not useful only in treating the mentally ill. It is important for all persons to have a healthy balance between work and play. Recreation offers the opportunity for relaxation, a change of pace, and time to pursue creative interests and develop meaningful and supportive social relationships with others. Particularly in an era in which so many persons lead isolated or alienated lives in an increasingly urbanized society, the value of recreation in self-discovery and as a form or creative personal release is crucial.

Main Idea: _____

Appendix B

Training Instructions for Anaphor Resolution Strategy

Adapted from Yuill & Oakhill (1988)

Introduction

When you say things or read things, sometimes there are short ways of saying them. For example, sometimes a person called David is called “Dave” for short. “Dave” stands for “David.” It is short for “David.” Sometimes in a story there are short ways of saying things. They are called anaphors. These anaphors always point back to something else in the story that has been said before, which are called referents. I am going to ask you about some of these anaphors. I will ask you to find the referent that the anaphor is referring to. Here are some examples:

1. Mary went for a walk. She found 10 cents on the ground.

   “She” stands for “Mary”, or points back to those words. You could put “Mary” instead of “she”, and the sentence would still mean exactly the same as it did before.

2. Mary went to the movies. John did too.

   You could say “John went to the movies too.” That is what “did” is short for, or points back to in the sentence. Both sentences mean the same thing.

3. “Is it raining?” asked Mary. “Yes,” said John.

   “Yes” stands for “Yes, it is raining.” Both mean the same thing. “Yes” points back to “it is raining.”

4. Mary was in her garden. The flowers smelled lovely.

   "The flowers" stands for "the flowers in her garden." The word "flowers" points back to the words "in her garden." The "flowers" mean "the flowers in the garden." Both mean the same thing.

Examples

Here are some more examples of this task. For example, for the sentences:

John went to the pool. He loved to swim there.

You would circle the word "pool" because it is the referent of the anaphor "there."

Try these 4 examples. For each one, circle the referent(s) that the anaphor is pointing back to.

1. On Saturday morning, Bill was going on a fishing trip with his uncle, the Captain. As he carried his fishing rod to the bus stop, he met Mrs. Tripp from next door.
2. "Who gave you your new fishing rod?," Alice asked. "Mom did," Bill replied.
3. "Did you remember the sandwiches?," asked Alice. "Yes," said Bill.
4. The Captain was trying to light his pipe. Each time, the wind blew out the tiny flame.

If perfect achievement is not achieved on the previous four examples, the following additional examples will be given.

Additional Examples

1. Bill smiled at Mrs. Tripp. He liked her because she often made a little cake just for him when she was baking.
2. “I hope you have a lovely time!” said Mrs. Tripp. “I hope I do too,” thought Bill to himself.
3. Alice had caught three fish. Poor Bill hadn’t got any.
4. Soon they were out on the open sea. The waves were high and the wind was blowing hard.

Appendix C

Eye Tracking Texts and Text Circling Measure

Adapted from Smith (2005)

Note: These text were adjusted for the counterbalanced research condition B in which the anaphor strategy was taught first.

Text 1

Please circle the words that you spent the most time looking at in this text.

Research is not a once-and-for-all-times job. Even sophisticated companies often waste its value. One of the most common errors is not providing a basis for comparisons. A company may research its market, find a need for a new advertising campaign, conduct her, then neglect to research the results.

Another may simply feel the need for a new campaign, conduct it, and research the results. Neither is getting its full benefit. When you fail to research either the results or your position prior to the campaign, you cannot know his effects. For evaluation you must have data before and after it.

Text 2

Please circle the main idea words in this text.

Until recently, the U.S. census, which is taken every ten years, offered only the following categories: Caucasian, Negro, Indian, and Oriental. It was a subject of much controversy because people felt that they did not allow for accurate identification of race. After years of their complaints, they were expanded.

In the year of 2000 census, everyone had to declare that they were or were not "Spanish/Hispanic/Latino." They had to mark "one or more races" that they "considered herself to be." Finally, if a person was still not happy with the choices, the person could check a box called "Some Other Race" and then write whatever was wanted. The new procedure allowed people to identify racial mix accurately.

Text 3

Please circle the main idea words in this text.

Prior to the time of Jan Baptiste van Helmont, a Belgian physician of the 17th century, it was commonly accepted that the source of plant matter was soil. (Probably, many people who haven't studied photosynthesis would go along with this today.) We aren't sure why, but van Helmont decided do an experiment. He carefully stripped a young willow sapling of all

surrounding soil, weighed the tree, and planted the tree in a tub of soil that had also been carefully weighed.

After five years of diligent watering (with rain water), van Helmont removed the sapling and again stripped away the soil and weighed him. It had gained 164 pounds. Upon weighing the soil, van Helmont was amazed to learn that it had lost only 2 ounces. She could not have lead to such growth in the tree. Plants seemed to be getting their matter from another source.

Text 4

Please circle the main idea words and underline the anaphor referents in this text.

If you're furious, you're at risk for an emotion charged confrontation. Such an interaction has an increased likelihood if you are experiencing that emotion. If you ambush someone with an angry attack, don't expect her to be in a productive frame of mind. If you need to resolve a conflict, give yourself time to relax before you try to address it. In the case of a group project, you could call a meeting for later in the week.

By that time, you could gain control of your feelings and better deal with one. Of course, sometimes issues need to be discussed immediately; you may not have the luxury to delay. But whenever it's practical, make sure you and your conflict partner are ready to communicate. Select a mutually acceptable time and place to discuss them.

Text 5

Please circle the main idea words and underline the anaphor referents in this text.

In one experiment, Daniel Lehrman of Rutgers University found that when a male blond ring dove was isolated from females, the dove soon began to bow and coo to a stuffed model of a female – one that it had previously ignored. When it was replaced by a rolled-up cloth, he began to court the new one; and when this was removed he directed one's attention to a corner of the cage, where he could at least focus its gaze.

The threshold for release of the behavior pattern became increasingly lower as time went by without the sight of the desired bird. Specific "energy" for performing instinctive courting behavior was building up within the male ring dove. Based on the findings, it was concluded that bird courting behavior is instinctive.

Text 6

Please circle the main idea words and underline the anaphor referents in this text.

In a Utah case, the defendant fell asleep in a car on the shoulder of the highway. Police stopped, smelled alcohol on his breath, and arrested her for driving while intoxicated. His conviction was reversed by the Utah Supreme Court because he was not in physical control of the vehicle at the time, as required by the law. In freeing the defendant, the Supreme Court judged that legal

sufficiency was not established in this case because the act observed by the police was not sufficient to confirm his guilt.

In other words, the case against the defendant failed because he was not violating the law at the time of his arrest and because it was also possible that she could have driven while sober, then pulled over, drank, and fell asleep. To establish legal sufficiency, he would have needed to be observed driving drunk.

Text 7

Please circle the main idea words and underline the anaphor referents in this text.

In recognition tests, retrieval cues (such as photographs) provide reminders of information (classmates' names) that we could not otherwise recall. Retrieval cues also guide us where to look for it. If we want to know what the pyramid on the back of a dollar bill signifies, we might look in Collier's Encyclopedia under "dollar," "currency," or "money." However, her words would not give it to us.

To find the information, we would have to look under "Great Seal of the United States." Like that stored in encyclopedias, memories are inaccessible unless we have good retrieval cues. They are better remembered when we have better methods of recalling them. Methods of creating and bringing up retrieval cues can be practiced to improve memory.

Appendix D

Strategy Use Survey

Note: These questions were adjusted for the counterbalanced research condition B in which the anaphor strategy was taught first.

Please answer each of the following questions. Circle the number that best describes your experience. Some questions ask how much you used a strategy and use the following scale:

1 = Not at all

2 = For one quarter or fewer of the relevant words

3 = For half or fewer of the relevant words

4 = For three quarters or fewer of the relevant words

5 = For all the relevant words

Other questions ask how useful you found a particular strategy and use the following scale:

1 = Not at all useful

2 = Somewhat useful

3 = Moderately useful

4 = Very useful

5 = The text would have been very difficult to read without using this strategy.

Please use the appropriate scale to answer each question.

First Text – The Elements of Good Evaluation in the Business World

1. How much did you use the anaphor strategy while reading this text?

   1 2 3 4 5

2. How useful did you find the anaphor strategy while reading this text?

   1 2 3 4 5

3. How much did you use the main idea strategy while reading this text?

   1 2 3 4 5

4. How useful did you find the main idea strategy while reading this text?

   1 2 3 4 5

Second Text – Changes in the U.S. Census Racial Categories

5. How much did you use the anaphor strategy while reading this text?

   1 2 3 4 5

6. How useful did you find the anaphor strategy while reading this text?

   1 2 3 4 5

7. How much did you use the main idea strategy while reading this text?

   1 2 3 4 5

8. How useful did you find the main idea strategy while reading this text?

1 2 3 4 5

Third Text – An Experiment on the Source of Plant Matter

9. How much did you use the anaphor strategy while reading this text?

1 2 3 4 5

10. How useful did you find the anaphor strategy while reading this text?

1 2 3 4 5

11. How much did you use the main idea strategy while reading this text?

1 2 3 4 5

12. How useful did you find the main idea strategy during while reading this text?

1 2 3 4 5

Fourth Text – Communicating During Conflict Situations

13. How much did you use the anaphor strategy while reading this text?

1 2 3 4 5

14. How useful did you find the anaphor strategy while reading this text?

1 2 3 4 5

15. How much did you use the main idea strategy while reading this text?

1 2 3 4 5

16. How useful did you find the main idea strategy while reading this text?

1 2 3 4 5

Fifth Text – Courting Behavior in Doves

17. How much did you use the anaphor strategy while reading this text?

1 2 3 4 5

18. How useful did you find the anaphor strategy while reading this text?

1 2 3 4 5

19. How much did you use the main idea strategy while reading this text?

1 2 3 4 5

20. How useful did you find the main idea strategy while reading this text?

1 2 3 4 5

Sixth Text – Legal Definition of Sufficiency in Drunk Driving Cases

21. How much did you use the anaphor strategy while reading this text?

1 2 3 4 5

22. How useful did you find the anaphor strategy while reading this text?

1 2 3 4 5

23. How much did you use the main idea strategy while reading this text?

1 2 3 4 5

24. How useful did you find the main idea strategy while reading this text?

1 2 3 4 5

Seventh Text – The Importance of Retrieval Cues in Memory

25. How much did you use the anaphor strategy while reading this text?

1 2 3 4 5

26. How useful did you find the anaphor strategy while reading this text?

1 2 3 4 5

27. How much did you use the main idea strategy while reading this text?

1 2 3 4 5

28. How useful did you find the main idea strategy while reading this text?

1 2 3 4 5

Appendix E

Metacognitive Reading Strategies Questionnaire (MRSQ)

Taraban, R., Kerr, M., & Rynearson, K. (2004)

Instructions: In this part of the survey, you will indicate what you do while reading. This part has 22 statements. Imagine that you are reading material for school. Take a moment to think about the typical things you do to help you comprehend the material. For each strategy statement, choose the statement that best indicates how much you use that strategy. Please read each statement carefully.

When information critical to my understanding of the text is not directly stated, I try to infer that information from the text.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

While reading, I write questions and notes in the margin in order to better understand the text.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

As I read along, I check whether I had anticipated the current information.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

I search out information relevant to my reading goals.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

I try to underline when reading in order to remember the information.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

I try to draw on my knowledge of the topic to help me understand what I am reading.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

I make notes when reading in order to remember the information.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

While reading, I underline and highlight important information in order to find it more easily later on.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

I read material more than once in order to remember the information.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

After I read a text, I consider other possible interpretations to determine whether I understood the text.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

While reading, I exploit my personal strengths in order to better understand the text. If I am a good reader, I focus on the text; if I am good with figures and diagrams, I focus on that information.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

I note how hard or easy a text is to read.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

As I am reading, I evaluate the text to determine whether it contributes to my knowledge / understanding of the subject.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

While I am reading, I reconsider and revise my prior questions about the topic, based on the text's content.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

While I am reading, I reconsider and revise my background knowledge about the topic, based on the text's content.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

While I am reading, I try to determine the meaning of unknown words that seem critical to the meaning of the text.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

As I am reading, I distinguish between information that I already know and new information.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

I evaluate whether what I am reading is relevant to my reading goals.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

While reading, I visualize descriptions in order to better understand the text.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

When I am having difficulty comprehending a text, I re-read the text.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

After I have read a text, I anticipate how I will use the knowledge that I have gained from reading the text.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

I anticipate information that will be presented later in the text.

I use this strategy

___Never

___Rarely

___Sometimes

___Often

___Always

Appendix F

Cloze Comprehension Questions

Note: These questions were adjusted for the counterbalanced research condition B in which the anaphor strategy was taught first.

First Set:

First Text – The Elements of Good Evaluation in the Business World

In order to effectively evaluate a campaign, one must compare data before a campaign to data _____ the campaign.

Second Text – Changes in the U.S. Census Racial Categories

The opportunity to mark several different racial categories on the census form allows individuals of more than one _____ to best describe their background.

Third Text – An Experiment on the Source of Plant Matter

We know today that photosynthesis is responsible for the growth of plants, rather than the belief that _____ is the source of plant matter.

Fourth Text – Communicating During Conflict Situations

Putting off a meeting until you are less stressed can _____ the risk of an emotion charged confrontation.

Fifth Text - Courting Behavior in Doves

When a male dove was isolated from female doves, it began to court _____ that resembled doves.

Sixth Text - Legal Definition of Sufficiency in Drunk Driving Cases

In the Utah Supreme Court case, legal sufficiency was not established because the defendant was not actually observed driving while _____.

Seventh Text – The Importance of Retrieval Cues in Memory

Unless we know which word to look up in the encyclopedia, we may not be able to find the _____ we are seeking.

Note: If 100% accuracy is not achieved, this second set was administered.

Second Set:

First Text – The Elements of Good Evaluation in the Business World

Some companies do not collect sufficient _____ to provide a basis for comparisons.

Second Text – Changes in the U.S. Census Racial Categories

People were upset about the U.S. census categories because they did not allow them to accurately identify their _____.

Third Text – An Experiment on the Source of Plant Matter

Van Helmont found that the _____ lost only a small amount of matter, while the tree gained a large amount of matter.

Fourth Text – Communicating During Conflict Situations

If you are stressed, you are _____ likely to resolve a conflict effectively.

Fifth Text - Courting Behavior in Doves

Over time, the substitute object needed to be less and less like a real _____ in order to elicit the courting behavior.

Sixth Text - Legal Definition of Sufficiency in Drunk Driving Cases

Legal sufficiency means that the act observed is _____ to establish guilt.

Seventh Text – The Importance of Retrieval Cues in Memory

Retrieval cues help us to look up information in encyclopedias, but also help us retrieve information from _____.

Appendix G

Pilot Studies

The texts for the experiment were piloted prior to beginning the experiment to ensure that the text were neither too difficult nor too easy, to ensure that the texts were at approximately equivalent reading levels, and to ensure that the participants comprehended the texts well.

The first pilot study attempted to determine if appropriate calibration could be achieved using the eye tracker. A second goal was to determine if there was a relationship between the words that participants spent the most time looking at in the texts (viewing a video of the participants' gaze during eye tracking) and the words that that participants said were most important in their understanding of the texts (self-reported by circling key words in printed versions of the texts). Four participants were enrolled in this pilot study. There was not a strong correlation between the gaze time for words on the screen and the participants' self-report of the importance of these words. Also, the words that participants found most important for their understanding varied greatly between participants.

Thus, a second pilot experiment was conducted to determine if providing metacognitive training to participants would affect their approaches to the texts, in that they would spend more time looking at the words involved in using the metacognitive strategies than at other words in the texts. Six participants were involved in this pilot study. Three received anaphor training and three received main idea training. Participants completed the reading task while monitored by eye tracking, and then circled the words in the texts that were relevant to use of the particular

strategy. In this pilot, the words participants circled in the texts did correlate well to the words that they spent the most time viewing during the eye tracking portion of the experiment. The gaze time to the words related to the strategy use was easy to measure because participants gazed at the words for a significant length of time and typically had some regressions back to these key words during reading of the texts. Several characteristics of the text were revealed in this pilot study which helped to create appropriate modifications to the texts. First, it was found that fewer of the anaphors were viewed than the main idea words. Second, main idea words were viewed more often when they were in groups of words, rather than single words. Modifications to the texts were made, based on these findings. First, two of the anaphors in each text were changed to be incongruent with their referent to increase the likelihood that they would be focused upon. Second, the main idea words were reorganized into continuous groups of words rather than single words throughout the texts.

With the validity of the metacognitive training and correlated increased gaze time for relevant key words for the two strategies established, the third pilot study determined the difficulty level of each text. This was determined through an analysis of the total reading time, number of comprehension questions accurately answered, and gaze time to the words involved in metacognitive strategy use. The DQW eye tracking program, which automatically calculates gaze time to different areas of the screen, was used to collect data for this pilot study. The location of each key word on the presentation screen was measured in each dimension and these measurements were converted from centimeters to pixel coordinates. Ten texts were read by eight participants and eye movements were monitored using the eye tracker. No metacognitive strategies were taught, and participants were asked only to read the texts silently. Participants

also answered comprehension questions about each text. The means and standard deviations for the total reading times for each text were generally consistent between texts (means=30.75-38.05 seconds, SDs=4.01-8.01 seconds), with the exception of one outlier (mean= 46.04 seconds, SD=9.15 seconds). This outlier text was removed from the set of texts used in the final experiment due to its long total reading time, relative to the other texts. Most participants answered each comprehension question correctly, with only 0-3 errors for each text. These questions were given to assure that the participants had a basic understanding of each text and were designed to be simple to answer if the participant had read the text. Finally, gaze times for the main idea and anaphor words were examined. There was a fair amount of variability in the means and standard deviations between the texts and between the two metacognitive strategies. This result was not unexpected because of the variability of gaze time noted during the first pilot study in which participants were asked to read texts without instructions about how to approach the text. Means ranged from 682.61 seconds to 3632.25 seconds, and standard deviations ranged from 376.76 seconds to 1367.59 seconds. Texts having the lowest mean (682.61 seconds for anaphor gaze time) and the highest mean (3632.25 seconds for main idea gaze time) were removed from the set of texts used in the final experiment, due to their inconsistency with the other texts. The final set of seven texts had a range of means from 885.10 seconds to 2909.97 seconds and a range of standard deviations from 376.76 seconds to 1367.59 seconds. Also, there was no significant difference between the gaze time for main idea words and anaphor words [$t(158)=-1.48$, $p=.14$] for these seven final texts.

Appendix H

Summed Gaze Times for Eye Tracking Texts

Table 1.
Summed Gaze Time in Milliseconds to Strategy Consistent, Strategy Inconsistent, and Control Words for Text 1

PARTICIPANT NUMBER	CONSISTENT TEXT 1	INCONSISTENT TEXT 1	CONTROL WORDS TEXT 1
2	1,987	684	2,438
3	935	1,436	2,889
4	2,070	551	3,373
5	801	751	1,920
6	2,204	1,419	2,872
7	417	1,169	2,705
8	1,619	2,454	1,302
9	1,837	2,672	2,070
10	818	434	5,678
13	384	1,603	2,538
15	751	501	2,087
16	267	1,068	2,020
17	901	651	3,707
18	2,037	2,939	1,135
19	1,503	985	4,976
20	1,887	2,839	6,062
21	901	2,020	2,588
22	935	233	851
23	283	1,419	3,423
24	1,820	2,154	2,989
25	467	367	2,221
26	384	467	2,555
27	1,052	951	4,575
28	2,087	167	2,989
29	601	3,540	4,258
30	3,573	5,110	4,074
31	551	2,221	4,124
32	985	1,970	918
33	2,171	2,037	3,406
34	1,252	801	1,903
35	434	4,676	3,690
36	3,340	283	2,889
37	3,089	2,772	2,755
38	1,753	1,352	5,093

Table 2.
Summed Gaze Time in Milliseconds to Strategy Consistent, Strategy Inconsistent, and Control Words for Text 2

PARTICIPANT NUMBER	CONSISTENT TEXT 2	INCONSISTENT TEXT 2	CONTROL WORDS TEXT 2
2	4,492	334	1,052
3	367	4,442	400
4	5,577	4,542	1,553
5	985	5,360	1,803
6	2,755	1,169	818
7	367	1,937	1,419
8	2,505	1,586	1,152
9	2,304	1,102	2,655
10	2,388	317	400
13	1,686	3,306	2,187
15	1,603	1,703	901
16	3,807	1,369	1,035
17	1,703	2,304	4,291
18	1,987	2,672	2,454
19	367	2,955	2,137
20	1,436	6,379	5,711
21	1,386	5,845	2,471
22	1,452	601	484
23	5,544	3,807	2,154
24	1,436	4,542	4,425
25	367	3,256	517
26	4,742	250	567
27	1,970	5,243	601
28	4,442	1,102	400
29	367	3,089	768
30	7,865	1,853	1,519
31	2,505	3,223	868
32	1,736	1,402	1,135
33	2,154	3,657	1,419
34	1,586	517	1,018
35	968	4,158	1,369
36	2,020	1,486	1,753
37	3,891	1,302	751
38	3,740	10,270	2,755

Table 3.
Summed Gaze Time in Milliseconds to Strategy Consistent, Strategy Inconsistent, and Control Words for Text 3

PARTICIPANT NUMBER	CONSISTENT TEXT 3	INCONSISTENT TEXT 3	CONTROL WORDS TEXT 3
2	1,085	4,826	4,575
3	4,008	2,488	5,394
4	968	2,371	2,555
5	6,863	1,619	4,759
6	1,486	5,678	6,663
7	3,507	901	3,757
8	384	3,139	4,041
9	3,941	1,486	3,139
10	3,690	6,012	6,212
13	1,953	567	2,839
15	3,373	2,287	3,640
16	1,920	2,905	3,740
17	4,158	718	7,464
18	1,018	5,911	3,423
19	3,740	2,972	6,362
20	200	10,754	2,237
21	3,423	1,586	3,273
22	784	2,471	2,388
23	4,943	1,569	4,759
24	1,920	7,047	2,070
25	1,753	1,770	6,730
26	1,302	4,258	5,895
27	2,705	1,519	7,631
28	1,686	5,410	7,915
29	3,740	1,152	7,297
30	6,763	10,153	7,698
31	6,128	1,670	6,362
32	1,185	5,193	2,037
33	7,798	634	8,550
34	1,035	3,807	2,588
35	2,388	1,686	4,826
36	1,235	3,573	7,114
37	4,909	317	3,256
38	2,020	5,611	12,825

Table 4.
Summed Gaze Time in Milliseconds to Strategy Consistent, Strategy Inconsistent, and Control Words for Text 4

PARTICIPANT	CONSISTENT TEXT 4	INCONSISTENT TEXT 4	CONTROL WORDS TEXT 4
2	4,275	668	1,302
3	2,104	3,790	3,807
4	1,219	3,891	1,018
5	7,214	2,505	651
6	3,373	3,039	1,603
7	6,379	501	350
8	1,853	1,052	784
9	3,189	3,273	450
10	4,525	1,920	3,941
13	2,722	2,538	501
15	7,999	1,469	2,421
16	2,605	2,889	1,336
17	734	5,511	3,924
18	3,106	534	668
19	7,849	501	668
20	6,963	784	1,219
21	1,770	3,991	350
22	2,354	1,068	2,438
23	4,024	6,346	3,824
24	1,786	3,256	1,118
25	4,876	4,175	3,540
26	1,786	5,260	1,369
27	9,436	501	1,870
28	2,571	4,041	2,237
29	2,438	3,323	5,143
30	2,605	2,872	3,707
31	9,435	1,102	1,436
32	601	4,475	3,156
33	3,122	1,736	567
34	3,373	7,815	1,937
35	9,436	3,790	4,659
36	1,035	2,271	668
37	634	4,726	4,843
38	2,054	11,973	2,070

Table 5.
Summed Gaze Time in Milliseconds to Strategy Consistent, Strategy Inconsistent, and Control Words for Text 5

PARTICIPANT NUMBER	CONSISTENT TEXT 5	INCONSISTENT TEXT 5	CONTROL WORDS TEXT 5
2	3,473	2,905	2,221
3	2,037	5,861	2,004
4	1,252	2,171	3,173
5	4,759	6,362	4,859
6	2,454	6,863	784
7	2,454	417	3,390
8	3,306	1,987	1,002
9	3,824	4,859	1,770
10	6,429	1,820	3,540
13	1,803	1,686	1,336
15	3,724	6,813	2,304
16	2,454	1,586	1,369
17	1,753	5,143	2,538
18	3,006	1,753	3,072
19	2,638	1,753	4,909
20	4,876	2,772	4,726
21	2,555	2,772	1,118
22	5,260	467	1,820
23	3,206	7,314	3,957
24	2,855	1,903	3,206
25	1,720	6,112	4,726
26	2,187	2,738	2,054
27	1,736	4,726	5,744
28	5,327	5,060	1,870
29	1,870	4,509	1,670
30	5,227	7,431	2,822
31	2,905	2,855	5,845
32	885	1,002	1,987
33	3,924	3,289	968
34	5,661	3,356	1,569
35	3,807	5,811	10,637
36	4,041	3,039	1,052
37	1,519	12,174	5,778
38	5,394	4,158	6,629

Table 6.
Summed Gaze Time in Milliseconds to Strategy Consistent, Strategy Inconsistent, and Control Words for Text 6

PARTICIPANT NUMBER	CONSISTENT TEXT 6	INCONSISTENT TEXT 6	CONTROL WORDS TEXT 6
2	2,070	4,291	701
3	7,932	718	701
4	1,219	718	3,456
5	3,757	1,987	1,636
6	1,786	3,056	3,640
7	183	1,202	1,920
8	1,887	3,206	300
9	6,162	901	1,569
10	5,761	2,705	4,208
13	1,185	1,452	2,154
15	1,152	1,786	701
16	3,523	2,404	1,152
17	8,734	2,588	718
18	1,720	2,955	1,169
19	183	1,736	3,223
20	7,214	4,926	1,035
21	3,273	1,603	818
22	1,436	534	150
23	7,732	818	668
24	1,636	3,707	2,287
25	3,423	1,135	250
26	2,137	2,404	1,503
27	501	2,438	3,523
28	6,412	3,072	868
29	5,444	885	835
30	4,659	1,736	1,085
31	183	2,388	2,004
32	2,571	1,553	2,354
33	2,621	1,436	2,037
34	2,054	2,104	1,536
35	183	3,690	2,976
36	3,273	1,887	1,386
37	7,615	3,774	150
38	4,208	2,555	150

Table 7.
Summed Gaze Time in Milliseconds to Strategy Consistent, Strategy Inconsistent, and Control Words for Text 7

PARTICIPANT NUMBER	CONSISTENT TEXT 7	INCONSISTENT TEXT 7	CONTROL WORDS TEXT 7
2	3,356	1,653	1,302
3	2,037	2,755	2,488
4	183	835	1,736
5	3,390	4,291	5,327
6	2,154	2,872	4,275
7	1,703	1,770	3,039
8	2,989	1,870	1,970
9	3,991	3,289	4,509
10	1,503	1,185	2,020
13	901	2,421	2,070
15	2,688	3,540	3,790
16	1,636	1,903	3,490
17	2,304	1,536	3,807
18	1,870	1,302	1,920
19	2,605	4,058	3,523
20	3,857	2,020	4,058
21	1,770	3,306	2,304
22	4,442	701	1,586
23	5,243	1,553	5,995
24	334	701	2,788
25	10,871	3,623	3,273
26	183	1,319	6,663
27	5,410	2,505	11,890
28	5,627	1,302	3,306
29	4,342	1,035	3,039
30	5,076	1,118	4,592
31	4,926	4,375	8,617
32	517	2,955	3,490
33	2,521	4,458	3,957
34	1,703	784	2,755
35	3,540	1,619	4,024
36	484	1,369	3,874
37	2,905	2,588	3,507
38	6,897	1,269	4,776

Appendix I

Calculation of Maintenance and Switch Ratios for Eye Tracking Texts

Table 1.
Computation of Maintenance Ratios Between Texts 2 and 3

PARTICIPANT NUMBER	CONSISTENT DIFFERENCE TEXTS 2 TO 3	CONTROL WORDS DIFFERENCE TEXTS 2 TO 3	MAINTENANCE RATIO TEXTS 2 TO 3
2	-3406	3523	0.97
3	3974	4993	0.8
4	-4609	1002	4.6
5	5878	2955	1.99
6	-1269	5845	0.22
7	3139	2338	1.34
8	-2120	2889	0.73
9	1636	484	3.38
10	1302	6212	0.21
13	267	651	0.41
15	1770	2738	0.65
16	-1887	2705	0.7
17	2454	3173	0.77
18	-968	968	1
19	3707	4225	0.88
20	-1235	-3473	0.36
21	2037	801	2.54
22	-668	1903	0.35
23	-601	2605	0.23
24	484	-2354	0.21
25	1720	6212	0.28
26	-3440	5327	0.65
27	734	7030	0.1
28	-2755	7849	0.35
29	3674	6529	0.56
30	-1102	6179	0.18
31	3623	5494	0.66
32	-551	901	0.61
33	5644	7130	0.79
34	-551	1569	0.35
35	1419	3456	0.41
36	-784	5360	0.15
37	1018	2505	0.41
38	-1720	10070	0.17

Table 2.
Computation of Maintenance Ratios Between Texts 4 and 5

PARTICIPANT NUMBER	CONSISTENT DIFFERENCE TEXTS 4 TO 5	CONTROL WORDS DIFFERENCE TEXTS 4 TO 5	MAINTENANCE RATIO TEXTS 4 TO 5
2	-801	918	0.87
3	-66	-1803	0.04
4	33	2154	0.02
5	-2454	4208	0.58
6	-918	-818	1.12
7	-3924	3390	1.16
8	1452	217	6.69
9	634	1319	0.48
10	1903	-400	4.75
13	-918	835	1.1
15	-4275	-116	36.57
16	-150	33	4.5
17	1018	-1386	0.73
18	-100	2404	0.04
19	-5210	4241	1.23
20	-2087	3507	0.6
21	784	768	1.02
22	2905	-617	4.7
23	-818	133	6.12
24	1068	2087	0.51
25	-3156	1185	2.66
26	400	684	0.59
27	-7699	3874	1.99
28	2755	-367	7.5
29	-567	-3473	0.16
30	2621	-885	2.96
31	-6529	4408	1.48
32	283	-1169	0.24
33	801	400	2
34	2287	-367	6.23
35	-5628	5978	0.94
36	3006	384	7.83
37	885	935	0.95
38	3340	4559	0.73

Table 3.
Mean Values of Maintenance Ratio Calculation

Block	Consistent Strategy Mean	Control Word Mean	Ratio
2	3040	2797	1.09
3	3121	2948	1.06

Table 4.
Switch Ratios for Texts 1-3

PARTICIPANT NUMBER	CONSISTENT/ INCONSISTENT TEXT 1	CONSISTENT/ INCONSISTENT TEXT 2	CONSISTENT/ INCONSISTENT TEXT 3
2	2.90	2.17	1.08
3	0.65	0.58	0.43
4	3.76	0.41	0.95
5	1.07	1.12	1.35
6	1.55	0.59	0.62
7	0.36	1.36	18.89
8	0.66	1.70	0.61
9	0.69	2.41	0.86
10	1.88	2.93	0.96
13	0.24	0.94	1.07
15	1.50	1.25	1.42
16	0.25	1.13	1.34
17	1.38	1.94	0.23
18	0.69	2.67	0.35
19	1.53	0.64	4.65
20	0.66	3.33	0.10
21	0.45	0.65	0.64
22	4.00	4.96	0.73
23	0.20	1.95	0.53
24	0.84	0.90	0.29
25	1.27	0.36	0.64
26	0.82	0.50	1.34
27	1.11	0.69	2.33
28	12.50	0.87	0.94
29	0.17	0.90	0.55
30	0.70	0.76	1.22
31	0.25	1.76	3.12
32	0.50	0.27	0.44
33	1.07	2.32	1.40
34	1.56	0.81	0.61
35	0.09	0.57	1.38
36	11.76	0.96	0.64
37	1.11	5.43	0.13
38	1.30	0.46	0.36

Table 5.
Switch Ratios for Texts 4-5

PARTICIPANT NUMBER	CONSISTENT/ INCONSISTENT TEXT 4	CONSISTENT/ INCONSISTENT TEXT 5
2	0.48	2.03
3	11.05	0.74
4	1.70	0.08
5	1.89	0.79
6	0.58	0.75
7	0.04	0.96
8	0.59	1.60
9	6.83	1.21
10	2.13	1.27
13	0.82	0.37
15	0.64	0.76
16	1.47	0.86
17	3.37	1.50
18	0.58	1.44
19	0.11	0.64
20	1.46	1.91
21	2.04	0.54
22	2.69	133.00
23	9.45	3.38
24	0.44	0.48
25	3.01	3.00
26	0.89	0.14
27	0.21	2.16
28	2.09	4.32
29	6.15	4.19
30	2.68	4.54
31	0.04	1.13
32	1.66	0.18
33	1.83	0.57
34	0.98	2.17
35	0.05	2.19
36	1.73	0.35
37	2.02	1.12
38	1.65	5.43

Appendix J

Reading Times for Eye Tracking Texts

Table 1.
Reading Times in Seconds for Texts 1-4 by Participant

PARTICIPANT NUMBER	READING TIME TEXT 1	READING TIME TEXT 2	READING TIME TEXT 3	READING TIME TEXT 4
2	32.03	36.92	53.73	41.05
3	30.37	46.13	47.58	43.68
4	35.68	46.48	40.43	35.10
5	21.67	48.95	57.70	66.58
6	38.83	45.37	68.97	56.55
7	34.75	33.80	37.02	40.20
8	27.73	29.17	38.30	24.78
9	34.93	37.42	45.75	42.62
10	31.70	39.29	77.75	39.00
13	33.15	37.68	37.23	37.55
15	28.00	33.53	49.37	48.22
16	23.90	37.91	54.21	35.88
17	36.57	47.13	65.16	55.18
18	32.27	34.87	53.94	34.37
19	32.78	39.67	60.61	56.16
20	61.95	65.63	73.82	45.72
21	31.07	33.42	45.58	43.50
22	14.78	20.85	26.79	21.32
23	30.30	64.98	55.88	69.76
24	39.61	45.52	52.48	44.58
25	38.45	43.45	60.45	70.25
26	40.52	46.66	57.94	47.05
27	35.50	64.94	56.20	69.51
28	25.75	71.12	81.11	46.85
29	39.22	29.47	57.75	39.32
30	63.53	69.32	102.42	47.27
31	37.72	43.23	58.75	55.17
32	30.84	30.96	41.28	32.10
33	44.18	43.54	76.01	47.27
34	23.17	41.73	48.17	38.74
35	41.85	45.75	60.77	136.63
36	35.35	42.60	57.07	34.90
37	47.33	48.63	59.72	56.80
38	54.97	82.20	130.03	64.52

Table 2.
Reading Times in Seconds for Texts 5-7 by Participant

PARTICIPANT NUMBER	RUN TIME TEXT 5	RUN TIME TEXT 6	RUN TIME TEXT 7
2	35.87	36.97	31.97
3	38.05	41.78	35.15
4	43.73	33.52	31.52
5	75.48	45.55	49.42
6	59.77	51.75	73.92
7	43.70	31.50	29.33
8	24.02	27.90	29.07
9	51.31	40.78	40.80
10	40.80	54.13	68.27
13	29.86	28.86	26.24
15	49.43	41.89	33.52
16	33.15	37.91	32.02
17	63.73	102.75	61.61
18	48.03	36.29	33.17
19	66.09	46.54	43.53
20	58.66	64.57	50.95
21	40.43	33.25	32.50
22	27.00	21.37	29.67
23	69.18	65.02	51.93
24	47.27	43.03	43.38
25	62.09	47.14	63.92
26	49.60	50.45	55.32
27	49.01	61.94	75.66
28	51.88	50.68	49.30
29	45.88	60.51	34.93
30	89.62	58.06	50.68
31	67.76	48.85	56.29
32	31.68	28.87	60.99
33	89.62	58.06	50.68
34	36.79	32.28	33.55
35	140.60	57.30	58.02
36	34.08	39.50	39.75
37	69.18	56.20	53.95
38	68.95	50.82	61.62