10-26-2009

Effects of Wide Reading Vs. Repeated Readings on Struggling College Readers' Comprehension Monitoring Skills

Omer Ari
Georgia State University

ACCEPTANCE

This dissertation, EFFECTS OF WIDE READING VS. REPEATED READINGS ON STRUGGLING COLLEGE READERS' COMPREHENSION MONITORING SKILLS, by OMER ARI, was prepared under the direction of the candidate's Dissertation Advisory Committee. It was accepted by the committee members in partial fulfillment of the requirements for the degree Doctor of Philosophy in the College of Education, Georgia State University.

The Dissertation Advisory Committee and the student's Department Chair, as representatives of the faculty, certify that this dissertation has met all standards of excellence and scholarship determined by the faculty. The dean of the College of Education concurs.

Gertrude Tinker Sachs, Ph.D.
Committee Chair

Brendan Calandra, Ph.D.
Committee Member

Joyce E. Many, Ph.D.
Committee Member

Rihana S. Williams, Ph.D.
Committee Member

Date

Dana L. Fox, Ph.D.
Department Chair, Middle-Secondary Education and Instructional Technology

R. W. Kamphaus, Ph.D.
Dean and Distinguished Research Professor
College of Education

AUTHOR'S STATEMENT

By presenting this dissertation as a partial fulfillment of the requirements for the advanced degree from Georgia State University, I agree that the library of Georgia State University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote, to copy from, or to publish this dissertation may be granted by the professor under whose direction it was written, by the College of Education's director of Graduate Studies and Research, or by me. Such quoting, copying, or publishing, must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without my written permission.

Omer Ari

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University library must be used in accordance with the stipulation prescribed by the author in the preceding statement. The author of this dissertation is:

Omer Ari
566 Laurel Creek Rd.
Sugar Grove, NC 28679

The director of this dissertation is:

Gertrude Tinker Sachs
Department of Middle-Secondary Education and Instructional Technology
College of Education
Georgia State University
Atlanta, GA 30303-3083

VITA

Omer Ari

ADDRESS: 566 Laurel Creek Rd.,
Sugar Grove, N.C. 28679

EDUCATION:

Ph.D.	2009	Georgia State University Teaching and Learning
M.A.	2003	Georgia State University Applied Linguistics and ESL
B.A.	2001	Bogazici University, Istanbul Foreign Language Education Department

PROFESSIONAL EXPERIENCE:

2009-2010	Clinical Instructor, The Reading Department, Appalachian State University, Boone, N.C.
2008-2009	Language and Literacy Doctoral Fellow
2008	Graduate Assistant at the Georgia State University Ronald E. McNair Post-Baccalaureate Achievement Program
2007-2008	Graduate Research Assistant at Tracking Eye-Movements and Vocabulary Acquisition (TEVA) Laboratory, Department of Psychology, GSU
2005-2008	Instructor of Regents' Reading Skills at MSIT, GSU
2003-2004	English as a Second Language (ESL) Teacher at Kaplan Educational Centers, Inc.
2003-2004	Teaching Assistant & Tutor, Intensive English Program, GSU
2001-2002	Tutor at ESL Tutoring Laboratory, GSU
2000-2001	English as a Foreign Language (EFL) Teacher at Istanbul Culture & Education Center, Istanbul

PROFESSIONAL AFFILIATIONS:

2008	National Reading Council
2006-now	International Reading Association
2005-now	Georgia Reading Organization

PRESENTATIONS AND PUBLICATIONS:

Williams, R. S., Ari, O., & Santamaria, C. N. (in press). Measuring college students' reading comprehension ability using cloze tests. *Journal of Research in Reading*.

Williams, R. S., Dortch, C., & Ari, O. (in press). African American high school students' reading comprehension performance as a function of enrichment program exposure. *Blacks in America: A Scholarly Response to the CNN Documentary*.

Ari, O. (2006). A review of three software programs designed to search for lexical bundles. *Language Learning and Technology, 10*, 1, 30-37.

Williams, R. S., Dortch, C., & Ari, O. (2008, October). *African American high school students' reading comprehension performance as a function of enrichment program exposure.* Paper Presented at the Think Tank Blacks in America: A Scholarly Response to the CNN Documentary, Tennessee.

Ari, O. (2008, April). College readers' fluency with print vs. speech discourse. Paper presented at Hard Data Café, Department of Psychology, Georgia State University, Atlanta, GA.

Santamaria, N. C., Ari, O., & Williams, R. S. (2008, March). *Measuring reading comprehension ability in college students using cloze tests*. Poster presented at the CEPO/Southern Psychological Association Conference, Charlotte, NC.

Wiliams, R. S, Ari, O., & Santamaria, C. N. (2008, February*). The use of cloze test as progress monitoring tools for college readers*. Paper presented at the 16th Annual Pacific Coast Research Conference in Coronado Springs, CA.

Davis, R., Ari, O., & Williams, R. (2008, February). *Using an open-ended cloze test to assess reading comprehension in adults*. Paper presented at the CLUES conference, Atlanta, GA.

Davis, R., Ari, O., & Williams, R. (2007, November). *Using an open-ended cloze test to assess reading comprehension in adults*. Poster presented at the Psychology Undergraduate Research Conference, Georgia State University, Atlanta, GA.

ABSTRACT

EFFECTS OF WIDE READING VS. REPEATED READINGS ON STRUGGLING COLLEGE READERS' COMPREHENSION MONITORING SKILLS
by
Omer Ari

Fluency instruction has had limited effects on reading comprehension relative to reading rate and prosodic reading (Dowhower, 1987; Herman, 1985; National Institute of Child Health and Human Development, 2000a). More specific components (i.e., error detection) of comprehension may yield larger effects through exposure to a wider range of materials than repeated readings (Kuhn, 2005b). Thirty-three students reading below college level were randomly assigned to a Repeated Readings (RR), a Wide Reading (WR), or a Vocabulary Study (VS) condition and received training in 9 sessions of 30 minutes in a Southeast community college. RR students read an instructional-level text consecutively four times before answering comprehension questions about it; WR students read four instructional-level texts each once and answered questions while the VS group studied and took a quiz on academic vocabulary. An additional 13 students reading at college level provided comparison data.

At pretest, all participants completed the Nelson Denny Reading Test, Test of Word Reading Efficiency, Error Detection task (Albrecht & O'Brien, 1993), working memory test, Metacognitive Awareness of Reading Strategies Inventory (MARSI; Mokhtari & Reichard, 2002), a maze test, Author Recognition Test (ART), and reading

survey. All pretest measures except for the ART and reading surveys were re-administered at posttest to training groups.

Paired-samples *t*-test analyses revealed (a) significant gains for the WR condition in vocabulary (p = .043), silent reading rate (p < .05), maze (p < .05) and working memory (p < .05) (b) significant gains for the RR students in silent reading rate (p = .05) and maze (p = .006) and (c) significant increases on vocabulary (p < .05), maze (p = .005), and MARSI (p < .005) for the VS group at posttest. Unreliable patterns of error detection were observed for all groups at pretest and post-test. Results suggest that effects of fluency instruction be sought at the local level processes of reading using the maze test, which reliably detected reading improvements from fluency instruction (RR, WR) and vocabulary study (VS) in only 9 sessions. With significant gains on more reading measures, the WR condition appears superior to the RR condition as a fluency program for struggling college readers. Combining the WR condition with vocabulary study may augment students' gains.

EFFECTS OF WIDE READING VS. REPEATED READINGS ON STRUGGLING
COLLEGE READERS' COMPREHENSION
MONITORING SKILLS

by
Omer Ari

A Dissertation

Presented in Partial Fulfillment of Requirements for the
Degree of
Doctor of Philosophy
in
Teaching and Learning
in
the Department of Middle, Secondary, and Instructional Technology
in
the College of Education
Georgia State University

Atlanta, Georgia
2009

ACKNOWLEDGEMENTS

I wish to thank the members of my Doctoral Advisory Committee for their support and assistance during the planning, implementation and presentation of this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF TERMS

Error Detection: A psycholinguistic task that is used to measure subjects' awareness of the state of their comprehension on narrative passages that are presented one line at a time on a computer screen. Passages are presented in a design that crosses the 2 levels (consistent vs. inconsistent) of the consistency condition with the 2 levels (local vs. global) of the coherence condition. Reaction times on the critical sentences are used in the analysis of the data.

Homograph: One of two or more words spelled alike but different in meaning or derivation or pronunciation (as the *bow* of a ship, a *bow* and arrow) (www.webster.com)

Maze: Maze is a progress monitor test in which every seventh word, except for the first sentence, is replaced with a 3-word choice. Subjects complete the maze test by circling the word that best completes the blank.

Reading Test: A high-stakes reading comprehension test which all undergraduates in the state are required to pass in order to graduate.

Reading Course: A course designed to give practice and teach strategies for the Reading Test.

Repeated Readings: An intervention program designed to provide practice to struggling readers through repeated readings of grade-level reading materials.

Vocabulary Study: An instructional program designed for study participants enrolled in Summer 2009 Reading Course. Reading Course students with grade-level reading skills (n = 3) and Reading Course students with below-grade level reading skills who were randomly assigned to serve as the control condition (n = 10) completed the procedures of the Vocabulary Study condition. The Vocabulary Study condition was designed because the research study took place during the class time of the students. Students in this condition studied 15 academic words and took two quizzes.

Wide Reading: An intervention program designed to provide reading practice to struggling readers through wider exposure to language. Students read the same amount of text as the Repeated Readings students but non-repetitively.

ABBREVIATIONS

CI	Construction-Integration
DV	Dependent Variable
ERDE	Error Detection
IV	Independent Variable
MARSI	Metacognitive Awareness of Reading Strategies Inventory
ND Reading Test	Nelson Denny Reading Test
RR	Repeated Readings
RSPAN	Reading Span
TOWRE SWE	The Test of Word Reading Efficiency: Sight Word Efficiency
TOWRE PDE	The Test of Word Reading Efficiency: Phonemic Decoding Efficiency
VS	Vocabulary Study
WM	Working Memory
WR	Wide Reading

CHAPTER ONE

INTRODUCTION

In their “simple” view of reading, Gough and colleagues dissociate reading ability into decoding and comprehension (Gough, Hoover, & Peterson, 1996). Viewed simply, reading is a product of decoding and comprehension; both are required, and neither is sufficient. Although the two are highly correlated in the general population, there may be variation in either of the components in more specific cases of dyslexics, hyperlexics, and garden variety readers (i.e., readers with low decoding and listening comprehension). They suggest therefore that “reading can be divided into two parts; that which is unique to reading, namely decoding, and that which is shared with auding, namely comprehension” (Gough et al., 1996, p. 2; see also Jackson & McClelland, 1979, who found listening comprehension and speed of letter matching to account for 77% of the variance in reading comprehension). In early grades, students are confronted with learning to recode written content to the auditory mode. At this stage, students are purposefully instructed through materials that they can aud well in order to facilitate the acquisition of decoding skills. Therefore, their reading comprehension depends on their success at recoding the written material to an intelligible mode (i.e., listening comprehension).

Over the years, however, as automatic word recognition skills develop, decoding ceases to determine reading ability; auding skills play a larger role in determining the success of their reading ability (Gough et al., 1996; Hoover & Gough, 1990b; Shankweiler et al., 1999). In other words, the strong decoding-reading ability relationship that characterizes early reading ability becomes negligible over time as students master the decoding skills and start reading content-rich reading material. At higher grades verbal ability (including knowledge of morphology, syntax, semantics, and pragmatics) starts to directly bear on reading comprehension ability (Hoover & Gough, 1990a; Juel, Griffith, & Gough, 1986). In other words, "given perfection in decoding, the quality of reading will depend entirely on the quality of the reader's comprehension; if a child's listening comprehension of text is poor, then his reading comprehension will be poor, no matter how good his decoding" (Juel et al., 1986, p. 244).

Students who grow up in literacy-restricted homes and do not develop rich vocabulary (Hart & Risley, 1995) are at risk for developing low verbal ability. They may perform equally well with a normative sample on all second- to third-grade reading and language tasks, but by fourth grade due to underdeveloped verbal skills their performance starts to decline, significantly on language tasks (e.g., the word meaning test in Chall, Jacobs, & Baldwin, 1990). Compensatory skills may help resist parallel declines in reading by sixth grade, after which the increasingly demanding content and vocabulary begins to depress students' reading comprehension by as much as 3-4 years (Snow, Barnes, Chandler, Goodman, & Hemphill, 1991).

Undoubtedly, access to print and motivation mediate students' interactions with reading (Guthrie, Wigfield, Metsala, & Cox, 1999; Neuman & Celano, 2001). However, their effects pale compared to decoding skills; with strong decoding skills, first graders are predicted to become avid readers who are exposed to a vast volume of print material (Cunningham & Stanovich, 1991, 1997, 1998; Stanovich & Cunningham, 1993; Stanovich, Cunningham, & Feeman, 1984). Print exposure in turn reinforces students' reading skills and builds their verbal ability. By reading avidly, not only do students develop greater cognitive skills and consolidate their knowledge of the alphabeticity principle, they are also exposed to more rare words and syntactic structures in print than in other media; children have 50% more chances of encountering rare words in children's books than in adult conversation or prime-time television shows (Hayes & Ahrens, 1988).

These findings urged authorities to push for content-based instruction and have found in Hirsch a vocal advocate (2003), who has developed school curricula featuring a set of core knowledge to accompany reading instruction. Verbal skills are regarded integral to developing fluency in reading; otherwise, students "may falter in their reading progress after making initially good progress" (Pikulski, 2006, p. 76). Without a commensurate focus on increasing students' access to and their engagement with challenging print material following the mastery of word recognition skills, poor comprehension may persist into college. Due to cumulative effects of scant exposure to print, college readers may not learn ordinary, general knowledge and disengage from reading. Lack of print exposure locks them in a downward spiral, leaving them with underdeveloped skills in the face of increasing academic demands and unprepared to understand material written for general public (e.g., a college reader with poor comprehension did not know about the Holocaust in a case study by Perfetti, Marron, & Foltz, 1996).

To get disenchanted students back into reading, several practice-inducing programs are available. Practice is assumed to increase struggling readers' motivation to read unfamiliar material with a greater sense of self-confidence and better word reading skills. Repeated readings or wide reading of connected text are the two major approaches to current fluency instruction. A greater emphasis is placed on students' word recognition skills in Repeated Readings, RR (Dahl, 1979; Samuels, 1979) in which students read a given passage either for a number of times or until they reach a preset criterion of words read correctly per minute. The RR program has proven to raise reading rate and word recognition skills of struggling students, with its effects for comprehension trailing (Dowhower, 1987; Herman, 1985; National Institute of Child Health and Human Development, 2000a). In Wide Reading (WR), on the other hand, students are provided opportunities to read material at their independent level and are exposed to a greater gamut of print language. Although highly recommended by many in the field, the program has failed to draw the blessings of the members of the National Reading Panel (National Institute of Child Health and Human Development, 2000b) due to inadequacies in statistical analyses reported by Wide Reading studies.

Rationale for the Study

In 1983, Allington alerted the literacy community that fluency was neglected in the nation's classrooms. Allington's call was well received; greater attention began to pour into fluency instruction and research (Adams, 1990; Snow, Burns, & Griffin, 1998). In 2000, the report of the National Reading Panel made its importance clear by adding it to four fundamental areas of reading instruction (i.e., phonemic awareness, phonics, vocabulary, comprehension instruction), and in 2002 it was incorporated into the No

Child Left Behind legislation. Yet, data from recent federal surveys once more alerted the literacy community to better understand the role played by fluency in reading and in remediating difficulties experienced by struggling readers.

The nation's report card, National Assessment of Educational Progress, has recently released data showing that reading scores of 12th-grade students are lower than a decade ago, so are those of 6th- and 8th-graders (National Assessment of Educational Progress, 2003). Another federal study, National Assessment of Adult Literacy (NAAL), has found that college graduates' basic reading scores are declining (National Center for Educational Statistics, 2001) despite a steady increase in the number of college enrollments, with 11% of the entering students needing remedial help in reading and writing. Disconcerting enough, the NAAL data state that today's college students are not graduating with basic reading skills: only 31 percent of college graduates are proficient (Romano, 2005) and only about a third of college graduates can read and understand a book. Prose proficiency was down at the time of the survey 10 percentage points for this group since the last assessment in 1992.

At a time when the field of cognitive psychology is making great strides in unraveling the workings of the human mind and demystifying the processes that underlie construction of mental models, we are encouraged to identify the true effects of fluency instruction that is starting to be widely used in the nation's classrooms (due to federal requirements such as No Child Left Behind Act, 2002) to promote fluency and comprehension scores of struggling readers. Using paradigms that are established in cognitive psychology, we are closer to achieving this goal.

Research in cognitive psychology and psycholinguistics have specifically revealed, for example, that less skilled readers are slower to execute higher order processes of integration for sentences (Gernsbacher, 1990; Gernsbacher, Varner, & Faust, 1990; Long, Oppy, & Seely, 1994; Long, Oppy, & Seely, 1997; Long, Seely, & Oppy, 1999) and for larger discourse which is artificially rendered challenging (Baker, 1985, 1989; Baker & Anderson, 1982; Cook, Halleran, & O'Brien, 1998; Long & Chong, 2001; Markman, 1979, 1981; Oakhill, Hartt, & Samols, 2005; Zabrucky & Moore, 1989). Agreed by most is the hypothesis that the underlying deficit in executing higher order skills rests in inefficient processing skills characterized by dysfluency (Daneman & Carpenter, 1980, 1983; Daneman & Merikle, 1996; Yuill, Oakhill, & Parkin, 1989). While good readers are able to detect inconsistent information that is embedded in passages, less skilled readers have shown large-scale insensitivity to such contradictions. The deficits in working memory are held responsible for less skilled readers' limited ability to represent inconsistent propositions and compare them (Vosniadou, Pearson, & Rogers, 1988). The effects are even greater when the working memory demands are increased with a manipulation of the distance between the inconsistent propositions: while skilled and less skilled readers are equally able to detect the inconsistent information in adjacent sentences, only skilled readers are able to do so in conditions where sentences containing inconsistencies are separated out (Long & Chong, 2001; Yuill et al., 1989).

Processing efficiency is viewed as a marker of poor comprehension by a group of researchers, who have pioneered the task-based processing theory of working memory (Daneman & Carpenter, 1980, 1983). Working memory is believed to underlie poor comprehension because of the simultaneous demands of multiple processes involved in reading, e.g., lexical access, parsing, integration, making inferences (Fletcher, 1994; Kintsch, 2004). While there are substantial skill differences on working memory tasks that require concurrent storage and processing (Yuill et al., 1989), skilled and less skilled readers do not differ on traditional, simple short-term memory tasks such as digit- or word-span (Oakhill, Yuill, & Parkin, 1986).

Therefore, research has surmised that due to underdeveloped processing skills poor readers are left with insufficient capacity to store text information. Readers with low working memory capacities are less likely to do well on standardized reading tests, and they are worse on more specific reading comprehension tasks, e.g., retrieving facts, detecting and recovering from internal inconsistencies, resolving anaphors separated from their antecedents by several sentences (Daneman & Carpenter, 1980, 1983). Furthermore, Daneman and Carpenter’s notion of working memory as task-specific has gained support from research findings that efficiency of symbolic (verbal and numerical) processes is a better predictor of reading comprehension than non-verbal, spatial span tasks (Daneman & Tardif, 1987; Shah & Miyake, 1996).

Fluency instruction, which was framed by the theoretical advances in the resurgent cognitive era of the 60s, was intended to improve the processing efficiency of struggling readers (Laberge & Samuels, 1974). Since its inception, fluency instruction has evolved in its implementation; however, its goal has remained largely unchanged. Research on the original application of fluency instruction, i.e., Repeated Readings, has substantiated the gains that are attributed to fluency instruction. For example, a number of studies and reviews have documented gains in the efficiency of poor readers’ oral reading fluency and comprehension (Dowhower, 1987; Herman, 1985; National Institute of Child Health and Human Development, 2000a; Rashotte & Torgesen, 1985). In addition, Herman (1985) and Dowhower (1987) reported increases in expressive reading characterized by longer phrases and fewer pauses.

However, a direct examination of the effects of fluency instruction on the less skilled readers' ability to execute higher order comprehension processes (integration) remains to be undertaken. Future research is warranted therefore to pursue this possibility given that there is a dearth of studies investigating specific effects of fluency programs in reading comprehension. We are further encouraged to examine as to whether the gains are due to repetitive practice or wider exposure to print, a debate stirred by recent research findings that wide reading may lead to greater gains in reading rate and comprehension than repeated readings (Kuhn, 2005a).

Purpose of the Study

The purpose of this study is to investigate the effects of two fluency intervention programs (Repeated Readings & Wide Reading) on maintaining global coherence in the texts read by poor comprehenders in college. Maintaining global coherence depends on efficient semantic and syntactic processing skills that ease the working memory constraints and allow readers to reinstate memory traces from distant portions of the text when they encounter an inconsistency. Informed by the theoretical insights of the Construction and Integration Model and by research in the fields of cognitive psychology and psycholinguistics (Albrecht & Myers, 1998; Albrecht & O'Brien, 1993; Kintsch, 1988), it is assumed that for comprehension to occur readers must preserve coherence in the mental models they construct. Constructing a coherent model is easier for students with good comprehension (Long & Chong, 2001; Long et al., 1994; Long et al., 1997; Long et al., 1999). In this study, the effects of fluency instruction on constructing coherent mental models were addressed in a group of college readers who have difficulty comprehending college level reading material due mostly to fluency deficits.

The study therefore (a) investigated whether fluency intervention leads to significant gains in constructing coherent mental models in a pretest-posttest research design (b) sought to specify the relationship between working memory and fluency, and (c) sought to identify the fluency program that leads to greater gains in helping struggling students construct coherent mental models.

Overview of the Study

The sample for this study comprised undergraduate students who read below the 13th-grade reading level on the *Nelson Denny Reading Test.* All undergraduate students, where this study took place, are required to pass the Reading Test in order to earn an undergraduate degree. Students can take the test only once a semester. The Reading Test, a high-stakes reading comprehension test similar to the *Nelson Denny Reading Test*, includes a total of nine passages and a total of 54 multiple-choice items to be completed in 60 minutes. Each passage is approximately 150 words long and is accompanied by six to eight comprehension questions. Among the questions are items assessing students' ability to retrieve verbatim information from the passage, to identify the meaning of unfamiliar words, and to draw inferences. The test consists of a higher proportion of inferential questions than questions that tap students' recall of verbatim information.

Data from the Board (2006), which oversees the administration of this test, indicate that only 33 percent of the repeating students passed the test in fall 2006. Of the students who took the test the first time in fall 2006, 78 per cent passed the test. These results suggest poor reading skills that persist in the repeating students. Previous research with repeating students substantiate this observation. Williams, Ari, and Santamaria (in print) found evidence that repeaters read on average at the 8.58th-grade level on the reading comprehension subtest of the *Nelson Denny Reading Test* compared to the 12.09th-grade level for the first-time students. On the vocabulary subtest of the *Nelson Denny Reading Test*, the difference between the repeating and fist-time students was 3 grade levels. The repeating students were also significantly behind the first-timers in the mean number of inference questions answered correctly on the reading comprehension subtest (F = 11.63, p < .05); while the first timers gave correct answers for a mean of 10.44 inferences questions, the repeating students answered only a mean of 6.82 inference questions correctly.

A fluency intervention program was implemented in Summer 2009 to remediate fluency deficits in a group of struggling undergraduate students enrolled in the Reading Course at a southeastern community college. Subjects reading below-college level were randomly placed in one of three conditions—Repeated Readings, Wide Reading, and Vocabulary Study—and completed the procedures of their respective condition for a maximum of nine sessions of 30 minutes in the span of three weeks. Students in the Vocabulary Study condition served as the control group and did not engage in any fluency-enhancing activity; their condition involved studying 15 academic vocabulary words, taking a quiz on the study words, and completing a Vocabulary Card for unfamiliar words per session. The Vocabulary Study condition was designed for the students randomly assigned to the control condition in order to keep them engaged during their class time while their peers in the Repeated Readings and Wide Reading conditions completed their fluency training.

The intervention was capped at 9 sessions on the basis of evidence from a pilot study conducted in Spring 2009 that fluency gains may be sufficiently attained in 8 sessions and that further training may not be necessary. Reading Efficiency scores (Reading Rate X Comprehension Accuracy) from 2 Repeated Readings and 4 Wide Reading students who completed a minimum of 16 sessions revealed the slope of gains in sessions 1-8 to be 1170.05 (t = 4.091, p = .006) and the slope of gains in sessions 9-16 to be 86.75 (t = .611, p = .564).

Fluency training involved reading grade-level passages drawn from *Timed Readings* (Spargo & Williston, 1975), a leveled series of books covering grades 4-13. Grade level achievement on pretest ND reading comprehension was used to match students to passages. In a Repeated Readings session, students read a grade-level passage four times back to back. On the other hand, Wide Reading students read four different passages per session each once. In addition to speed, students were cued to read for comprehension. Ten comprehension questions were answered by Repeated Readings students after the fourth reading of the passage; three questions per passage were answered by the Wide Reading group totaling 12 questions per session. Records of reading rate and comprehension were kept from each session.

Data collection commenced at pretest when paper and pencil measures (Nelson Denny Reading Test, maze, MARSI, Reading Survey) as well as computer-based tests

(ERDE, RSPAN, TOWRE) were administered. Weekly maze tasks were administered to provide a measure of progress monitoring. Gains on reading comprehension, vocabulary, silent reading rate, MARSI, word recognition, working memory, and error detection were determined from a comparison of the posttest data with the pretest data. Various Univariate Analyses of Variance were performed to analyze the data.

CHAPTER TWO

THEORETICAL FRAMEWORK

Because the goal of this research project is to investigate the impact on higher-order comprehension skills of underachieving readers of fluency instruction, the proposed study is appropriately framed in the Construction-Integration (CI) Model (Kintsch & van Dijk, 1978), which will be discussed below. The discussion will continue with working memory—an essential component of comprehension and a central source of individual differences in reading comprehension—and its role in the construction of mental models.

The Construction-Integration (CI) Model

According to the CI model, there are three layers of discourse representation: (a) surface, (b) textbase, and (c) situation model (Kintsch, 1988, 2004). The first layer consists of the text as it appears on the page in words, sentences and paragraphs. However, language input from text is not represented in its surface form; a semantic representation is required. Words are represented in relation to other words in one's lexicon by their features or category. They may also be represented on dimensions of the semantic space (e.g., Latent Semantic Analysis; Landauer & Dumais, 1997). Sentences are thought to be represented as idea units that are stripped off the surface structure. Idea units, which are also called propositions, construct a relation between words in a sentence (or a phrase); for instance, for the sentence *He cooked imambayildi last night* a relation between the agent, the object and adverb of the sentence is constructed around the predicate *cooked*. Because the underlying semantic form is used for the cognitive process of comprehension, propositions are not affected by syntactic forms; the same proposition is constructed for a sentence in passive or active voice. In other words, propositions provide a language for the cognitive theorist to investigate the mental representation of the discourse (Kintsch, 2004; Perfetti & Britt, 1995).

Psychological reality of propositions has been verified in a priming study by Ratcliff & Mckoon (1978), in which the subjects were faster to recognize a target word taken from the same proposition as the prime than to respond to a target word primed by a word from a different proposition. The distance between a prime and a target in the surface form did not change the observed reaction data: words that were adjacent but belonged to different propositions were slower to prime each other than words that belonged to the same proposition.

A cyclical construction process is conceptualized for transforming the surface structure into a network of propositional representations, or a textbase. This process moves through the text in cycles of sentences or phrases, creates propositions out of these linguistic units in conjunction with the comprehender's prior knowledge, and maps each text input to the previously processed propositions that are maintained in the short-term memory buffer. A referential or a causal relationship guides the mapping process whereby propositions that share a concept are interconnected and establishes local coherence in the developing mental model. If propositions in the cycle do not share a concept, a search is initiated in all previously processed propositions. If the search finds a proposition that shares a relation with the current input in working memory, the input is mapped onto the developing model. If the search fails to find a matching concept, then an inference is generated that connects the current input to the propositions in the textbase. Therefore, a processing cycle includes the current sentence that enters the working memory and the immediately preceding two to three propositions as well as "important" (i.e., tightly-connected) propositions. The size of a cycle, however, is assumed to depend on reader and text characteristics; for a difficult or unfamiliar text, the input size might be reduced to a smaller set of phrases or sentences (the boundaries of input that enter the working memory) than a familiar text. Familiarity also affects the capacity to store propositions from the earlier sections of the text as the reader's mental resources are devoted to analytical processes of decoding, syntactic and semantic integration, and making inferences. Unlike the difficult texts, readers execute these processes in easier texts without much consciousness since most of the lower level processes (e.g., decoding, syntactic and semantic integrations) are executed automatically. Skill in reading similarly affects the amount of information that enters the cycle and the number of propositions retained from the earlier cycles (Kintsch & van Dijk, 1978).

In a more connectionist version of the model revealed by Kintsch (1988), textual input activates multiple mental representations of words and propositions in the comprehender's background knowledge for each cycle. While some of the activated associations may be relevant, some of them may be irrelevant, or even contradictory because of the careless and "dumb" memory-based activation. Therefore a cycle may be characterized by inconsistency until integrative processes enforce consistency in the model through a connectionist procedure called spreading activation, which strengthens the appropriate outcomes and reduces the activation of irrelevant constructions. For example, multiple meanings of a word may be activated at the first encounter and only after a short delay (i.e., 350ms) the meaning that is most constrained by the context is maintained (e.g., the money and riverbank meanings of the word BANK, Swinney, 1979). Integration within a cycle is therefore a function of the context-appropriate activation of the words and propositions. Activation within a cycle may also be determined by the previously constructed propositions that relate to the topic of the text. In other words, gist information carried over from earlier portions of the text reinforce a globally-appropriate interpretation of the propositions and reduce an irrelevant outcome within each cycle (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983).

The predictions of the model were supported in a study by Till, Mross, and Kintsch (1988), in which college students were presented sentences such as *The townspeople were amazed to find that all the buildings had collapsed except the mint*. Following each sentence, target words were presented in a lexical decision task; students were asked to press a *yes* key if the target was a word and press *no* if it was not. The target words were either an associate of the final word (money vs. candy) or a topical inference word (earthquake vs. an unrelated word). In addition, the target words were presented at different time intervals ranging from 200 ms to 1500 ms. Both appropriate (money) and inappropriate (candy) associate words were activated initially (at 200ms-300 ms) but only the appropriate associate (money) remained active on the following intervals; the inappropriate associate was deactivated after the 300-ms interval.

This finding illustrates that all associated meanings of words are activated at the first encounter and that only the appropriate sense is selected by contextual constraints.

The inappropriate sense is dropped out of the textbase for the sentence. A similar pattern of activation was observed for the topical inference word (earthquake) relative to an unrelated topical control word. However, the integrative process for the topical inference was not completed until as long as 500ms after the offset of the sentence. Both related and unrelated topical words were active in the first 500ms following the sentence. After 500ms, the unrelated topical inference word was deactivated and finally dropped from the mental model.

As the local level construction is under way, the construction-integration process generates a discourse level representation that not only provides contextual constraints and reduces inconsistency within a cycle but also links the current production to the overall theme of the discourse. For such a global representation to be derived from the discourse, van Dijk and Kintsch (1983) list three macro rules that act upon the propositions or sequences of propositions in the microstructure (i.e., deletion, generalization, and construction). Propositions that do not share any arguments (unrelated to the rest of the text) are deleted from the developing textbase. Propositions that can be subsumed under a general proposition are replaced with a general one. Or, a series of propositions may lead to the creation of a related proposition (i.e., inferring “paying” from “shopping”, Kintsch & van Dijk, 1978, p. 366). Applied to the developing textbase, these rules reduce the textbase to a summary of propositions called a macrostructure consisting of major gist information in the form of macropropositions. This ongoing process of deleting, generalizing, and constructing is carried out until the entire text is condensed to one summary proposition. Studies using priming and recalling found evidence that subjects are faster at responding to a macroproposition when it is primed by a macroproposition than by a microproposition and that subjects recall more propositions from a macrostructure than a microstructure at delays of up to three months (Kintsch & van Dijk, 1978).

A coherent textbase is required for successful comprehension. However, a textbase representation (including the micro and macro structure) of discourse alone is not sufficient for the message derived to be useful in the future and cannot be applied to new situations naturally when needed. The textbase must be integrated within the general knowledge of the reader for learning to occur. If the textbase is not rooted in prior knowledge, it may be maintained as a capsule of information, which can only be retrieved when the reader is reminded of the text from which it was derived (Kintsch, 2004). Readers' goals and interests play an important part in the integration of the textbase into his/her world knowledge and yield an idiosyncratic situation model which may involve imagery in addition to the propositional representations derived from the textbase. Two readers may construct the same textbase but may differ in the interpretations they draw from the same textbase because of their different backgrounds (Fletcher, 1994; Glenberg & Langston, 1992; Graesser, Singer, & Trabasso, 1994; McKoon & Ratcliff, 1992; Morrow, Greenspan, & Bower, 1987).

Coherence in Developing Mental Representations

Coherence reflects the extent to which the mental model is well connected around the main theme of a discourse. Absent coherence, readers may construct mental models that are choppy and disjointed. Gernsbacher and colleagues argue, for example, that lack of coherence may lead readers to initiate a new substructure when the incoming text does not cohere with the recently comprehended information (Gernsbacher et al., 1990). It becomes difficult to map incoming input to the developing mental structure; readers therefore branch out a new substructure if they cannot suppress the irrelevant information. If not suppressed, irrelevant information lays the foundation of a new structure and results in poor access to information in the previous structure. The effect is greater for less skilled readers.

Although most reading comprehension models account for coherence, there is disagreement among them as to the level of coherence that is routinely maintained. In a minimalist model for example, built on the cycle-by-cycle construction process of the CI model in which inferences are generated to fill gaps between propositions, local coherence is the main concern of the comprehension process. In this model, a reader generates inferences when s/he encounters a break in local coherence; readers normally generate bridging inferences that construct an unstated relation between sentences, for example. The reader does not encode other inferences unless other, general information is “easily available” during the course of reading (McKoon & Ratcliff, 1992, p. 441). A constructionist model, however, rejects the minimalist view, by arguing that the number of inferences generated by readers under normal circumstances far exceeds a “minimal” number. While the “minimalist” reader encodes inferences that are ordinarily generated in the absence of reader goals, a “constructionist” reader, in contrast, employs strategies even during ordinary text processing, by engaging in a “search after meaning” (Graesser et al., 1994, p. 371), whereby s/he makes inferences that explain why actions, events, and states are mentioned in the text, reads in a way that addresses the goals s/he initially sets and attempts to construct a mental model that is coherent both at local and global levels. The debate over reader goals is not so much about the validity of the memory-based processing as the role of reader goals during discourse comprehension: the constructionist model institutes goals during the ordinary memory processes; however, the minimalist view rejects this assumption.

Recent research has provided evidence for memory-based processing that transcends the debate. O’Brien and colleagues have shown that without a conscious effort on the part of the reader, information from a distant part of the text becomes passively available, even when a text is locally coherent and thus there is no need for strategic search of information. Inactive information becomes available through a *resonance* process in which the current content of working memory activates relevant portions of the long term memory, including the inactive portions of developing discourse representation and the general world knowledge. In response to the signal from the working memory, concepts from long-term memory resonate with the information in the working memory as a function of their match to the input. The memory traces that signal back to the input in turn resonate with other traces residing in the long term memory. As the resonance proceeds, the most active traces from long term memory (that share most semantic and contextual overlap with the input in working memory) reenter working memory (Dell, McKoon, & Ratcliff, 1983).

A word in the introduction of a short passage, for instance, may drop out of working memory after it is backgrounded by a section of unrelated text and may regain accessibility at a point where a related context is rementioned (Gerrig & O'Brien, 2005). A series of experiments by O’Brien and colleagues (O'Brien, Rizzella, Albrecht, & Halleran, 1998) has also indicated that memory activation is *dumb*; earlier information that was introduced as inconsistent but was later changed to be consistent was still activated at the point where students read the target sentence. For example, in one of the short passages, Mary was introduced to be a strict vegetarian, which causes an inconsistency at the target sentence which states that she orders a cheeseburger and fries. The inconsistency causes readers to experience a comprehension difficulty, which is manifested in their slower reading times. However, the reading times are still slow on the target sentence (that she orders a cheeseburger and fries) after a qualified elaboration that Mary is not a vegetarian anymore.

In discussing the dumbness of the memory based processing, Gerrig and O'Brien note that "memory processes cannot assess truth value: Even though the inconsistent characteristic was not true [in the case of Mary], the information continued to be activated and affected comprehension" (2005, p. 232). Further research has also shown that reactivation is unrestricted (Cook et al., 1998; Long & Chong, 2001); even when focus is changed from Mary to a secondary character who is introduced to be the vegetarian, at the target sentence where Mary orders a cheeseburger and fries, a probe task reactivated characteristics of the secondary character at the target sentence. A plausible explanation is that the secondary character was reactivated because his characteristic (of being the vegetarian) shared features in common with Mary's action of ordering a cheeseburger and fries. A goal-directed explanation could not account for such a finding because the target sentence is both locally and globally coherent: Mary who is not a vegetarian orders a cheeseburger.

To understand a memory-based processing view of comprehension (constructing mental models), a closer examination of the working memory mechanisms are necessary. In the following section, working memory will be defined, its relation to reading and reading comprehension will be delineated and research implicating working memory deficits in reading comprehension differences will be discussed.

Working Memory as a Domain-Specific System

Reading comprehension is a complex behavior for which skilful execution of multiple cognitive processes is required both at lower and higher levels (Long, Johns, & Morris, 2006). As postulated by dominant theories of reading ability (Perfetti, 1985) and comprehension (Just & Carpenter, 1992; Kintsch, 1974, 1988), inefficient processing at lower levels could cause a bottleneck for constructing a coherent mental model and executing higher order skills involved. For example, Perfetti's Verbal Efficiency Theory (VET) predicts that slow and inaccurate word identification will inevitably consume attentional resources that are needed for higher level comprehension skills (e.g., monitoring comprehension, making accurate inferences, and so on). In this theory, deficiencies at word recognition are presumed to tax the reader's working memory and stifle efforts to construct a coherent model of the text. What ensues is a discussion of working memory theories and the role working memory plays in constructing coherent mental models.

After the publication of the Baddeley & Hitch (1974) chapter, the concept of working memory has taken hold among cognitive psychologists. This new conception of working memory as a short-duration, limited-capacity system with storage and computation functions was a marked departure from the earlier conceptions of memory as storage space and entry point to the long-term memory. In fact, the storage component has come to be viewed by some as *superfluous*; Daneman and Tardif (1987) suggest storage component of working memory is wholly dependent on the processing component. This suggestion was based on equally strong correlations between verbal ability and a verbal working memory span task with and without the storage component in the studies Daneman and Tardif conducted.

In Baddeley and Hitch's model (1974), the executive system is the center for processing while the slave systems of articulatory loop and the visuo-spatial sketchpad are delegated with storing the processed information. The slave systems are specialized for storing specific kinds of material; the articulatory loop manages the verbal stimuli while the sketchpad is tuned into storing the visuo-spatial stimuli. Conclusive evidence corroborates the notion of specialization for the slave systems. In several studies with children and adults verbal and math span tasks predicted variance in reading comprehension, but a spatial task failed to predict any variance (Baddeley, Logie, Nimmo-Smith, & Brereton, 1985; Nation, Adams, Bowyer-Crane, & Snowling, 1999; Olesen, Westerberg, & Klingberg, 2004), suggesting that working memory as it relates to reading "is specialized for manipulating and representing symbolic information" (Daneman & Tardif, 1987, p. 500).

Furthermore, STM and WM seem to be two distinguishable constructs in studies of factor analysis which investigated the underlying construct(s) to a set of tasks thought of tapping the short-term memory capacity versus those thought to be involved in working memory performance (Cantor, Engle, & Hamilton, 1991; Engle, Tuholski, Laughlin, & Conway, 1999).

This more functional, task-specific concept of working memory has been supported in studies contrasting skilled and less readers on simple storage and newer storage plus processing tasks of working memory (Oakhill, 1984). Simple storage tasks did not distinguish skilled readers from less skilled readers; however, the more demanding storage plus processing tasks have consistently predicted skill difference on complex cognitive tasks (i.e., reading comprehension). Poor comprehenders seem to have normal digit span and verbatim recall and do not differ from normal readers in short-term memory (Oakhill et al., 1986; Perfetti & Lesgold, 1977; Stothard & Hulme, 1992).

Perhaps the most direct assessment of task-specific view of working memory in reading comprehension was provided by the hallmark Daneman and Carpenter (1980) study. To investigate the sources of differences in reading, a reading span task was developed by these researchers who believed that simple word or digit span tasks did not account for complex reading processes involved in reading. Storage of the final or intermediate products is conditional on the processing efficiency of the readers. Readers who require fewer processes or who are not required to process intermediate processes (in decoding, lexical access, parsing, inferencing, integrating) would have a greater capacity to store the outcome of the processes during comprehension. This trade-off between processing and storage is what distinguishes skilled from less skilled readers and is not captured by storage only tasks of memory (Perfetti & Goldman, 1976).

Daneman and Carpenter (1980) found that while the reading span task was related to college students' Verbal SAT (r=.59), it showed greater association with the more specific comprehension tests of fact retrieval (.72) and answering pronoun-reference questions (.90). Readers with smaller spans were worse off on both comprehension tasks than readers with larger spans. While the low span readers answered only 5.4 of the 12 pronoun reference questions, the high span readers answered 9.7 questions. In contrast, a word span task (storage only) was significantly less related to fact and referent retrieval. On the pronoun reference questions, in which the distance between the pronoun and its referent was varied from two to seven sentences, an interaction between reading span and distance was found. Students with the span size of 3 had difficulty finding the right referent for the pronoun over 2-3 sentences. For span 4 readers, the difficulty arose when the distance was over 5 sentences. Span 5 readers showed no errors in finding the right referent for the pronoun even for sentences at a distance of 6-7 sentences.

Computing noun-pronoun relations and recalling facts require the reader to hold in memory the products of earlier processes while processing the current information and relating it to the earlier discourse segments. Unlike the complex working memory tasks, simple memory tasks (e.g., a word span task in which the only task requirement is to recall as many of the words as presented, Perfetti & Goldman, 1974), fail to differentiate good from poor readers on general as well as specific reading tests with greater demands of processing and storage.

Furthermore, consistent with the task specific view of working memory, researchers (Lee, Lu, & Ko, 2007) have found that training in a visuo-spatial task (mental abacus) improves participants' visuo-spatial storage and processing capacity, confirming the notions of working memory as an experience-based system (Ericsson & Kintsch, 1995). The fact that the training effect was limited to a visual-spatial task and did not extend to tasks that tap verbal working memory (digit span and non-word span) supports domain-specific arguments of working memory (Daneman & Tardif, 1987; Shah & Miyake, 1996).

Similarly, increasing the rate at which sentences are read seems to overcome the constraints of the limited capacity and rapid decay associated with Short-term Memory (STM). Breznitz and Share (1992) have shown that when the per-letter presentation rate of sentences is adapted to the highest per-letter reading rate a reader is capable of sustaining, performance improves significantly on the STM-sensitive tasks as well as on tests of reading comprehension, decoding accuracy and rate. Second-graders' performance on the STM-sensitive tasks used in this study suggests that fast-paced reading induces greater STM encoding of the text. In a series of experiments, Breznitz and Share observed large gains in second graders' word and order recall in addition to a recency effect, all markers of STM processes. The specificity of the effect was clearly demonstrated in a task of detecting wording or semantic changes to short passages. In the detection task, a passage was presented followed by its unaltered original version and one of two altered versions: a version in which a content word was replaced with a synonym and a version in which a nonsynonymous change was made. As opposed to self-paced condition, the fast-paced condition produced larger detection of the original sentences when the students distinguished between the original version and the version with a wording change. In other words, students reading at a fast rate retained the wording of the original passages during the detection task. An opposite pattern was observed under the self-paced reading condition, suggesting a tendency for meaning retention during self-paced reading and rapid decay of memory for exact wording.

The task-based model of working memory has, however, been challenged by Engle and colleagues who have posited controlled attention as the cause to individual differences observed in working memory tasks (Conway & Engle, 1996; Engle, Cantor, & Carullo, 1992). In a moving window presentation of both the operation span and reading span, Engle and colleagues were able to control for the effects of the processing component in the working memory tasks involved. The correlation between the span scores derived from the WM tasks and the reading comprehension test was not affected when the processing time was partialled out. The WM-higher order cognition relationship was shown to be unaffected by Conway and Engle in an Operation Span task taken by college students. Participants' processing ability was obtained on a pretest measure arithmetic task, data from which were used to adapt the difficulty of the processing component of the operation span task for each participant. This manipulation of the difficulty of the processing component ensured that participants performed at 75-95% correct levels on the operation component. The correlation between performance on the adjusted operation span task and reading comprehension was not different from the one in which an unadjusted operation span task was used.

The more processing efficiency differences are implicated in differential reading comprehension, the more salience is ascribed to print exposure and practice as a precursor to the deficits in the processing skills of poor comprehenders. What follows is a description of the evolution of the theories from general capacity resources to skill-through-experience, connectionist theories of working memory.

Theories of Working Memory

Despite theoretical differences, recent models of working memory have concurred with Baddeley and Hitch's (1974) functional notion of working memory (i.e., processing and storage). Most theories based their assumptions on findings from studies of syntactic ambiguity resolution by readers of low and high working memory span. For example, King and Just (1991) had college students read relative clause sentences such as *1a* and *1b* below and recorded their reading times on the main verb *admitted*.

1a. The reporter that attacked the senator admitted the error.

1b. The reporter that the senator attacked admitted the error.

Analysis yielded three significant effects: relative clause type, span, and the interaction of both on the main verb. The reading times were shorter in the subject relative clauses (sentence a) than the object relative clauses (sentence 1b) when the data were collapsed over reading span groups. The reading times were shorter in the high span group when the reading times were collapsed for the relative clause type. On the easier subject clause type (1a), both groups had similar reading times of the main verb *admitted*, but on the harder object clause (1b), low span readers exhibited lower reading times.

Interpreting reading time data from King and Just (1991), Just and Carpenter (1992) argued that the difference in the reading times in the harder object relative clause (sentence 1b above) between the high and low span participants is due to working memory capacity. All readers had enough capacity to comprehend the easier subject relative clause (sentence a above) but it was only the high span participants who had sufficient capacity to read the harder object relative clause (sentence 1b above).

The Capacity As Activation Model was supported in studies of reading times on sentences that included verbs with either a Main Verb or a Reduced Relative interpretation (MacDonald, Just, & Carpenter, 1992). In this research with college readers, low- vs. high-span college undergraduates (based on the reading span task of Daneman & Carpenter, 1980) read sentences with ambiguous verbs. The verbs denoted either a Main Verb interpretation or a Reduced Relative interpretation such as the following:

2a. Main Verb Resolution—Unambiguous: The experienced soldiers spoke about the dangers before the midnight raid.

2b. Main Verb Resolution—Temporarily ambiguous: The experienced soldiers warned about the dangers before the midnight raid.

2c. Relative Clause Resolution—Unambiguous: The experienced soldiers who were told about the dangers conducted the midnight raid.

2d. Relative Clause Resolution—Temporarily ambiguous: The experienced soldiers warned about the dangers conducted the midnight raid. (p. 61)

The verb *warned* is an ambiguous verb with two possible interpretations: (a) as the past tense form of the verb it may denote an action that the subject did or (b) as the past participle form of the verb it may be used in reduced relative clauses in which the relative pronoun that*/who/which* and auxiliary verb *was/were* are omitted. Reading a sentence with this verb leads to ambiguity as to which interpretation needs to be chosen.

MacDonald and colleagues (1992) hypothesized that high span readers can keep both interpretations in memory until they reach the disambiguation region at the end of the sentence. Low span readers, however, due to capacity limitations are unable to hold two representations at the same time; they drop the "unpreferred" representation from the memory and hold only the preferred representation. Low and high span readers are therefore expected to show different patterns of reading time at the end of the sentence where they encounter the disambiguating word. The high span readers, because they kept two representations in memory at the expense of processing some information starting from the ambiguous region, "wrap up" at the end of the sentence and complete the unfinished processing job at that region, now that there is only one interpretation left and the least likely interpretation can be dropped from their memory. The low span readers show no wrap-up effects at the end of the sentence because they did not hold more than one interpretation in memory. They chose only one representation and processed the sentence through that representation. When they were at the last word, there was no leftover processing to do for them.

Therefore, in their Capacity Constraint Parsing Model, MacDonald and colleagues (1992) argue that a high-span reader maintains more than one representation in memory when she encounters an ambiguity. Starting at the ambiguous region, the reader postpones most of the language processing in order to keep the second representation in memory. The words that s/he buffered (stored in memory unprocessed) are processed when s/he reaches the disambiguating region. Consequently, the reader who can hold two representations in memory takes longer on the disambiguating word than a reader who does not have such capacity.

A different conceptualization of working capacity was offered by Waters and Caplan (1996) called Separate Sentence Interpretation Resource (SSIR), also using relative clauses like the ones below. In this theory, syntactic processing is modular and does not vary across individuals. A second part of the memory system acts upon the sentence meaning following the construction of a syntactic representation (SSIR) and is the locus of individual differences: making inferences to integrate ideas across sentences, drawing on world knowledge, and so on.

3a. Subject relative clause—The scout warmed the cabin that contained the firewood.

3b. Object relative clause—The cabin that the scout warmed contained the firewood (Caplan, Waters, & Dede, 2007, p. 273).

Using both online (reading times) and offline tasks (responding at post-processing), Caplan and colleagues found that normal readers exhibit slower reading times on the working memory demanding region of object relative sentences and are slower to judge the acceptability of object relative sentences, compared to subject relative sentences, in an offline task (Waters, Caplan, & Hildebrandt, 1987). The difference in the online tasks is thought to represent the time it takes to integrate the words presented one at a time into the developing mental representation of the sentence. For the regions with higher working memory demands, in the object relative sentences, the integration process is slower. The amount of information that must be integrated and the distance over which such processing needs to take place determines the reading time; in object relative clauses, the embedded clause requires storage of information from sections of the sentence (distance) and processing of the embedded clause (computation or processing requirement). The longer listening times at the verb of the object relative clause confirm the differential working memory demands by object and subject relative sentences; in the subject relative sentences, the corresponding regions (the clause final word) had relatively shorter latencies.

Waters and Caplan (2004) found support for their hypothesis in a college sample who were tested on working memory capacity using a variant of the Daneman and Carpenter sentence span task they developed (plausibility judgment) and on syntactic processing. In an auditory moving windows task, sentences such as the following (subject and object relative sentences) were presented as students listened to each sentence phrase by phrase while their listening times were being recorded:

Subject Relative

4a. It was/ the book/ that/ interested/ the teenager.

4b. The millionaire/ favored/ the law/ that / frustrated/ the workers.

Object Relative

4c. It was/ the teenager/ that/ the book/ interested.

4d. The law/ that/ the millionaire/ favored/ frustrated/ the workers (Caplan et al., 2007, p. 275).

In the plausibility judgment task, participants were asked to judge if the sentences were acceptable or not before recalling the final word; the task was developed as a more reliable alternative to the Daneman and Carpenter reading span task (Waters & Caplan, 1996b). Subjects were classified as low, medium, or high WM span based on their performance on the plausibility judgment reading span task. The results revealed that listening times for all three groups (low, medium, high span subjects) were higher on the phrases of the object relative clauses which exerted greater capacity demands: the verbs *interested* and *favored* in the object relative clause above, relative to corresponding positions in the subject relative clauses.

Alternative approaches to the relationship between capacity and reading comprehension include skill- and experience-based views (Ericsson & Kintsch, 1995; MacDonald & Christiansen, 2002). MacDonald and Christiansen's connectionist theory frames WM variation in differential language experience and is supported by successful simulations of the reading time data reported by King and Just (1991) and Just and Carpenter (1992). MacDonald & Christiansen's simulated models comprise a network of processing units which differ in number and weights of activation; model's structure and experience determine the capacity of the system, not a separate pool of capacity resources as in WM theories of Just and Carpenter (1992), Just and King (1991), and Waters and Caplan (1996).

For example, in explaining the successful replication of King and Just (1991) data on subject and object relative clauses, MacDonald and Christiansen argue that subject relative sentences have a regular word order (like simple active one-clause sentences) and are very frequent in English. Object relatives, on the other hand, have irregular structures and are rather infrequent in English. Experience with simple sentences aid both high and low span subjects to comprehend subject relative clauses but may not be sufficient to comprehend rare object relative clauses, for which direct experience with object relative sentences is required. MacDonald and Christiansen further argue that Span X Clause Type interaction found in King and Just (1991) is phenotypic of the Frequency X Regularity effect found in word recognition by Seidenberg (1985) and Pearlmutter and MacDonald (1995).

Integration of Information and Working Memory

In discourse comprehension, working memory is believed to underlie a reader's ability to integrate semantic and syntactic representations in a sentence and to construct coherent discourse-level representations. The comprehension processes that distinguished high- from low-span readers in the classic Daneman and Carpenter (the pronoun reference task and recalling facts, 1980) study are a major source of individual differences in reading-related processes (i.e., decoding, lexical access, parsing, integration, inferencing), implicating deficits in these processes. What follows is a summary of the previous research in constructing mental representations for sentences and larger discourse separately.

Sentence-Level Integration

Research in sentential processing adapted the Till, Mross, and Kintsch (1988) paradigm to examine the differences between skilled and less skilled readers. In the Till et al study, students completed a lexical decision task following the presentation of a sentence that ended with a prime. Among the targets in the lexical decision task were associates of the prime word (the sentence final word) and a topical inference word for the sentence. The integrative process of spreading activation resulted in the deactivation of inappropriate associates after a delay of 350ms following the sentence offset. The delay in making a topical inference took readers 150ms longer than the selection of the appropriate associate. That is, the readers required as long as 500 ms to deactivate an unrelated topical inference word relative to a related inference word. These findings however served to establish a precedent for the psycholinguistic research in sentence processing in a homogeneous group of college readers. Unlike this study, the following studies examined the effects of delayed sentence processing of skilled readers in comparison to less skilled readers, with a conclusion that the delayed processing observed in less skilled readers are due to the processing deficits and working memory capacity limitations.

Long and colleagues (1997) formed skilled and less skilled groups from a sample of college undergraduate students based on their Verbal SAT scores and adapted the design and materials of Till, Mross, and Kinstch (1988). The subjects were to read 2-sentence passages, study them, and press YES if a test word appeared in the passage and press NO if it did not in the testing session that ensued. The results indicated that both groups were faster to recognize (in a priming paradigm adapted from Ratcliff & McKoon, 1978) test words from the same proposition than a different proposition as the prime. Likewise, they responded faster to the appropriate sense of the ambiguous homograph (e.g., *mint* with two meanings, see Table 2 below) than to the inappropriate sense; there was no difference between the groups in response times.

However, the groups differed in the topic condition: Not only was the skilled group slower to reject appropriate topic words when they saw the prime, their error rate was also higher than that of the less skilled readers, who showed no difference between appropriate and inappropriate topic words in both reading times and error rates. The following table displays the design of the priming paradigm and the topic-related target words that were primed by associates of the homograph *mint*.

Table 1

Example Passages and Target Items from Long et al. (1994)

Passage	Target items	
	Prime	*Topic*
The townspeople were amazed to find that all the buildings had collapsed except the *mint.* Obviously, it had been built to withstand natural disasters.	money	earthquake
Thinking of the amount of garlic in his dinner, the guest asked for a *mint.* He soon felt more comfortable socializing with the others.	candy	breath

Slower response times and higher error rates by the skilled readers were taken as an indication that skilled readers made topic-related inferences. They might have failed to reject appropriate-topic words (thus higher error rate) because they had already included these topics in their mental representations through accurate inferences. Or, as was suggested by the authors, they made backward associations between the topic word and the mental representation of the passage during the test phase. It appears that the difference in making topic-related inferences is limited to online processing; when skilled and less skilled readers were tested on their knowledge of the passages offline, both groups make correct topic-related inferences (Long et al., 1994; Long et al., 1997).

These findings suggest that poor comprehenders show structural awareness of propositions but are slower to construct a semantic representation for the sentence online. Long and colleagues (Long et al., 1997) argue that less skilled readers do not seem to have a deficit in making inferences. Their inference problems appear to emerge *during* comprehension, when integrative processes are deployed to reinforce activated memory nodes that are appropriate and deactivate the inappropriate ones. This argument was based on the fact that less skilled readers made correct topic-based inferences (a task that requires integration) when they were asked for a word that described the situation in a two-sentence passage after carefully studying it.

Delayed construction of sentence representation of less skilled readers can also be observed in a meaning fit judgment task, which taps students' ability to suppress irrelevant associates of a homograph (Gernsbacher et al., 1990). In a meaning-match experiment, skilled and less skilled college readers verified if a test word following a sentence matched the meaning/context of the sentence. For example, in sentence 5a, the last word *spade* is an ambiguous word and is an associate of *ace*. However, 5b ends with *shovel*, an unambiguous word which shares no meaning association with the test word *ace*. For both examples, the subjects are to press the NO key because the test word *ace* does not fit the meaning of the sentences.

5a. He dug with spade.

ace

5b. He dug with shovel.

ace

Immediately after reading the test word, the skilled comprehenders experienced a significant amount of interference. The inappropriate meanings were still highly active at the immediate interval. The interference that the skilled readers demonstrated at the immediate interval diminished significantly at the delay, however; the inappropriate meaning became considerably less activated about a second later. Skilled readers were able to suppress the inappropriate meaning at the delay. Less skilled comprehenders also experienced a significant, similar amount of interference at the immediate interval. However, the less skilled readers were still experiencing a significant amount of interference after the delay; almost the same amount they experienced at the immediate interval.

The results indicate that the groups differed in how fast they finished the integrative process for the sentence. While the skilled group finished this process after the immediate interval (200ms), the less skilled group was still engaged in the process at least at the delay condition, about a second later. Because both groups were still processing the sentence at the immediate interval, they were vulnerable to the interference from an associate word. However, the associate ceased to interfere for the skilled readers at the delay who had by this time finished processing the sentence and freed their working memory. Less skilled readers on the other hand never finished constructing a coherent sentence representation, even until the delay, and therefore were still vulnerable to the interference caused by the associate *ace* of the sentence-final word *spade*.

Long, Seely, and Oppy (1999) later showed that suppression deficits are due to processing inefficiency that arises in tasks requiring integration or context checking, i.e., meaning-fit judgment that was used in Gernsbacher and colleagues' study (1990). Both skilled and less skilled readers were found to be similarly able to suppress inappropriate associate of a homograph in a naming task, which is immune to context checking; however, the less skilled readers experienced a great deal of interference from the inappropriate associate when the task was a lexical decision or meaning-fit judgment task. In particular, less skilled readers were influenced by the context in making the judgments in the lexical naming task whereby they were to indicate if the test probe was a word or not and in the meaning-fit judgment task in which they were to indicate if the test probe was related to the meaning of the context, such as the following:

6a. The townspeople were amazed to find that all the buildings had collapsed except the mint. *Money*

6b. Thinking of the amount of garlic in his dinner, the guest asked for a mint. *Candy (p. 300).*

At the immediate condition, for both groups the inappropriate meanings were activated; the subjects were still processing the sentence and thus devoting their mental resources to this process. However, at the delay it was only the less skilled readers who still had inappropriate associates activated, probably because they were still trying to construct a mental representation for the sentence and did not have sufficient attentional resources to inhibit the response conflict introduced by the relation between the sentence-final word and its inappropriate associate. It appears from these study results that sentential semantic and syntactic integrative processes take longer for unskilled readers who may construct structurally adequate propositional representations (i.e., show priming effect to a word from a word from the same proposition than a different proposition; Long et al., 1997).

Further evidence for delayed integrative abilities of less skilled readers can be found in a study by Long, Oppy, and Seely (1994), who showed that less skilled readers may not even start the integrative process at the end of the sentence and therefore be able to complete a lexical decision task immediately after reading the sentence, unaffected by sentence reading. However, they are slower on a second lexical decision task at delay (200ms after the offset of the sentence) because they start processing an integrative meaning for the sentence at this point and become susceptible to working memory limitations. An opposite pattern of reading times was observed for skilled readers: they do not suffer from a 200-ms lag effect since they start end-of-sentence processing earlier.

They are slower on a second lexical decision task immediately after reading the sentence, but they recover from the lag effect after the first 200ms (when they finish the integrative process) and outperform the less skilled readers on the lexical decision task (Long et al., 1999). To rule out the possibility that the lag effect observed in this study was due to something else than integrative processes, Long et al. (1994) replicated the experiment with scrambled passages; the lag effect was eliminated. Since the need to integrate words is obviated in the scrambled condition, the groups did not differ. The skilled group did not show the lag effect as early as they did and the less skilled groups did not suffer from the lag effect as late as they did.

Discourse-Level Integration

Comprehension of a sentence in discourse is more than accessing the meanings of its words and integrating the meanings most activated. It must be integrated with preceding information in the discourse. This requires reactivation of earlier segments of the text while the current sentence is being processed. Memory-based models of comprehension have suggested an automatic resonance process in which earlier information is reactivated by the information currently in working memory. A signal is sent out to the long-term memory including discourse representation and the general world knowledge of the reader. Concepts or ideas that share semantic and contextual overlap resonate and enter the working memory. This resonance process has been documented even when the text is locally coherent and there is no need for reactivating long-term memory (Albrecht & Myers, 1995; Albrecht & O'Brien, 1993; McKoon, Gerrig, & Greene, 1996; Myers, O'Brien, Albrecht, & Mason, 1994). The reactivated information in turn is integrated with information currently being processed.

According to Myers and O'Brien (1998), in addition to the automatic, dumb resonance process, a second process evaluates the integration between reactivated information and the current content of working memory. This second process may flag the reader about an integration failure and may cause the reader to engage in problem solving (van den Broek, Risden, & Husebye-Hartmann, 1995). This may happen when the resonance process fails to yield an appropriate antecedent for an anaphoric reference, when the automatic process fails to identify an appropriate candidate from among a number of distracters, when the process fails to activate an antecedent cause for the current consequence, or when there is a contradiction between reactivated memory traces and propositions in the focus of working memory. Operationally, strategic engagement in integration can be observed in longer reading latencies. Research in discourse comprehension has confirmed these assumptions in a series of studies using reading time data. Slower reading times were found on sentences containing an anaphor or a causal consequent for which the antecedents were sufficiently backgrounded (O'Brien, Plewes, & Albrecht, 1990; Rizzella & O'Brien, 1996). Slower reading times were also found for sentences that contradicted earlier portions of a text relative to those which were consistent (Albrecht & Myers, 1995; Albrecht & O'Brien, 1993; Myers et al., 1994). For example, in Albrecht and O'Brien (1993) college students read a passage about a protagonist; the passage was divided into three sections: introduction, elaboration, and conclusion. In the conclusion section, subjects' reading times were recorded on the critical sentence *As Bill was talking to Mrs. Jones, he saw a young boy who was lying in* *the street hurt. He quickly ran and picked up the boy over to the curb…* The elaboration of the protagonist as having just celebrated his twenty-fifth birthday and was feeling in top condition was consistent with this critical sentence. However, the elaboration that the protagonist had just turned eighty-one and did not feel as strong as he was twenty years ago was inconsistent.

It took the subjects longer to read the critical sentence in the inconsistent condition than the consistent condition, supporting the resonance hypothesis. These results indicate that normal college readers experienced comprehension difficulty on the critical sentence in the inconsistent condition, even though the inconsistent characteristic (that Bill had just celebrated his eighty-first birthday and did not feel as strong as twenty years ago) was removed from the working memory by a filler section of several sentences. In accordance with the resonance model, the physical characteristics of the protagonist were reinstated when the readers encountered the target sentence in which he performs a physically demanding action for an eighty-one-year-old. These findings support the resonance model and explain the role of memory-based processes in the construction of locally and globally coherent mental models.

Although these findings confirm the theoretical assumptions of the memory-based processing, they do not provide empirical data about memory-based processes in less skilled readers' discourse comprehension. In a comparative study of good and poor college readers, Long and Chong (2001) adapted Albrecht and O'Brien's (1993) paradigm and included a distance variable (global coherence vs. local coherence) in the passages describing two characters (e.g., Ken and Mike, see Appendix A for a sample passage). The filler section between the introduction and the target sentence was reduced from 77.6 words in the global coherence condition to 15.1 for the local coherence condition. After two characters were introduced (Ken and Mike), one of the characters was elaborated by either a short filler (15.1 words) or a long filler (77.6 words). Following the elaboration the target action was presented performed always by the first character (Ken).

An inconsistency effect was observed for good readers both at the local and global conditions for the relevant character only (Ken). Poor comprehenders in contrast exhibited an inconsistency effect only at the local condition for both characters. In other words, although the poor comprehenders were able to detect the inconsistency in the local condition, they failed to differentiate between the characters.

A probe-verification task was presented to test the possibility that poor comprehenders failed to detect the inconsistency in the global condition because the character elaboration was backgrounded by a long filler section. Probe sentences about the characters were presented either after the character description, before the target action, or after the target action. Both good and poor comprehenders were equally faster to respond to the probe after the description section and the target sentence. Consistent with the Resonance model, both first and second-character elaborations were activated at the target sentence, even though only the first character description was relevant to the action. This finding rejects the hypothesis that poor comprehenders fail to maintain global coherence because they lose access to information earlier in the text. Poor comprehenders had no difficulty reinstating prior text information into memory and "character description was sufficiently well encoded in poor comprehenders' text representations to be reactivated by information in the target action" (Long & Chong, 2001, p. 1428).

Long and Chong (2001) argue that poor comprehenders' difficulty is at the discourse level, where integration is required between propositions even when each proposition may be encoded well enough to be reactivated from long-term memory. In this sense, their findings are consistent with other studies showing that poor comprehenders experience difficulty constructing globally coherent text representations (Cain, Oakhill, Hulme, & Joshi, 1998; Garnham, Oakhill, & Johnson-Laird, 1982; Long et al., 1994; Long et al., 1997; Oakhill & Yuill, 1986).

Further evidence that working memory deficits underlie construction of coherent mental models is offered by a line of research on bridging inferences, known also as text-based or backward inferences (Singer, 1993; Singer, Andrusiak, Reisdorf, & Black, 1992; Singer & Ferreira, 1983). Text based inferences are known to accompany comprehension and are generated to fill in a logical gap (Singer & Ferreira, 1983; van Dijk & Kintsch, 1983). For example, in the following sentence the reader has to generate a causal bridge to make sense of the passage; the two sentences without the bridging causal relation that the fire burned the report cannot be mapped onto a coherent representation.

7a. The spy quickly threw his report in the fire. The ashes floated up the chimney (Singer & Ritchot, 1996, p. 733)

The reader utilizes the idea that fire burns something like a report made of paper, adding a new proposition in his/her mental representation of this short text. Without the bridge, an accurate mental representation cannot be derived. To test this hypothesis, Potts and colleagues (Potts, Keenan, & Golding, 1988) had subjects read the following sentences and the probe word *Dead*. Reading latencies for the probe word were recorded and compared across conditions 8a-8c.

8a. The director and the cameraman were preparing to shoot closeups of the actress on the edge of the roof of the 14^{th} story building when suddenly the actress fell.
The director was talking to the cameraman and did not see what happened.
Dead

8b. The director and the cameraman were preparing to shoot closeups of the actress on the edge of the roof of the 14^{th} story building when suddenly the director fell over the camera stand.
The director was talking to the cameraman and did not see what happened.
Dead

8c. The director and the cameraman were preparing to shoot closeups of the actress on the edge of the roof of the 14^{th} story building when suddenly the director fell over the camera stand.
Her orphaned daughters sued the director and the studio for negligence.
Dead

Students' response times to 8c was slower than to 8b, indicating that students made a backward inference from the second sentence in 8c "Her orphaned daughters sued the director and the studio for negligence" to the first sentence when they saw the probe word (dead) and therefore they were slower to read the probe in this condition. This however was not the case for sets 8a and 8b.

The processing and storage requirements involved in generating accurate bridging inferences illustrate the substantial demands placed on the working memory capacity of the reader: the reader must construct a mental representation for the first sentence, hold that representation in memory, do the same for the second sentence, relate the two sentence representations to each other, create a bridge that adequately accounts for this relationship, and validate the newly created bridge against world knowledge by accessing prior knowledge (Singer & Ritchot, 1996). For example, in the following sets, Singer and Halldorson (1996) found that readers were faster to answer the question in 9c after reading 9a than after reading 9b.

9a. Dorothy poured the bucket of water on the bonfire. The bonfire went out (causal).

9b. Dorothy placed the bucket of water by the bonfire. The bonfire went out (temporal).

9c. Does water extinguish fire?

To validate the causal relation in 9a, the reader has to access the mediating idea that water extinguishes fire. For readers who are slow to construct mental representations for the two sentences, the verification time on 9c is longer (Singer et al., 1992; Singer & Ritchot, 1996). In a more direct test of working memory constraints, Singer and Ritchot formed groups of high/low span and high/low access from a sample of college readers. Daneman and Carpenter's reading span test was used to form the span groups, while

Potts and Peterson's (1985) integration task was used to form the access groups in their ability to access prior knowledge. The students read passages in which the bridging sentences were presented either adjacently (near) or were separated by four intervening sentences (far). The passages were presented one sentence at a time on a computer screen. An example of far and near conditions are presented below:

10a. Near:

1. Valerie left early for the birthday party (motive)
2. Valerie left the birthday party early (control)
3. She spent an hour shopping at the mall.
4. Do birthday parties involve presents?

10b. Far:

1. Valerie left early for the birthday party (motive)
2. Valerie left the birthday party early (control)
3. She checked the contents of her purse.
4. She backed out of the driveway.
5. She headed north on the freeway.
6. She exited the Antelope Drive.
7. She spent an hour shopping at the mall.
8. Do birthday parties involve presents? (p. 736)

A significant main effect of working memory span was found: high span readers answered the questions (d or h) 107ms faster than did low span subjects. By contrast, a high span high access group was 77ms slower on the control passages compared to the motive passages. The authors attribute this “counterintuitive” finding to Just and Carpenter’s (1992) argument that high span readers have greater working memory capacity and therefore can hold multiple interpretations in memory. On the other hand, a low span low access group had the slowest answer times on the motive condition (i.e., when there was strong impetus for making a bridging inference); they had a difference of only 81ms from the control condition. The authors argue that these subjects labor at word recognition and proposition construction—the lower level language processes—and take too long to validate the bridge against their world knowledge. Without a sound mental model constructed for both sentences and kept in memory, these subjects may not be able invoke world knowledge in creating a bridge between the two sentences.

The role of working memory was also directly researched later by Hannon and Daneman (2001) in their “theoretically sound” and practical new measure of Component Processes Task. The task is an improved version of Potts and Peterson’s (1985) integration task. The new task accounts for a greater amount of variance (about 60%) in a college group’s reading comprehension than most of the single-resource tasks used previously (e.g., working memory).

The new task taps four components in a single measure that lasts only 30 min. The components include accessing prior knowledge, integrating accessed knowledge with text information, making inferences based on text information, and recalling new text information from long-term memory (Hannon & Daneman, 2001, p. 121). In one of their experiments designed to validate the new measure and compare it with other components of reading comprehension (e.g., working memory), Hannon and Daneman found that of the four components, accessing prior knowledge failed to correlate with working memory; text inferencing and knowledge integration were significantly correlated with working memory than accessing prior knowledge (correlations at around .50). Working memory was also found to share no variance with speed of accessing prior knowledge. It appears from these findings that working memory shares variance with such higher order processes as making text based inferences and integrating information from long term memory with information derived from the text in reading comprehension, but not with accessing prior knowledge, nor with the speed of access.

Summary

Research evidence suggests that poor readers fail to execute integrative processes in sentences and larger texts as fast as good readers (Garnham et al., 1982; Long & Chong, 2001; Long et al., 1997; Oakhill et al., 2005; Seigneuric, Ehrlick, Oakhill, & Yuill, 2000; Yuill et al., 1989). Poor readers integrate words in sentences at a lower rate than good readers and therefore show a pattern of delayed construction of sentence representations (Long et al., 1994; Long et al., 1999). Furthermore, poor readers' difficulty in integrating text information seems to be exacerbated by increasing working memory demands. Poor readers fail to show insensitivity to inconsistencies spread out in the passages while they may be able to detect the inconsistencies in adjacent sentences. Therefore, text representations of unskilled readers are sketchy and disjointed. Unskilled readers can remember single propositions relying on their memory, but they are unable to integrate propositions to construct a holistic unit to represent the discourse (Oakhill et al., 2005; Seigneuric et al., 2000). Most attribute poor readers' failure to construct coherent mental models to inefficient processing skills: processing word meanings while executing syntactic and semantic integrative processes for sentences and relating them to a discourse-wide representation (Daneman & Carpenter, 1980, 1983).

The review of the literature on working memory suggests processing to be task specific and subject to improvement as a result of experiential factors (e.g., practice and training) in the specific domain (visual vs. verbal) (Daneman & Tardif, 1987; Lee et al., 2007). Recent findings in working memory and its interaction with reading ability urge us to think that less skilled readers may be trained to increase their processing skills through fluency programs designed to provide practice (e.g., Repeated Readings). In the next section, I will review the literature on training struggling readers. Two widely used programs, Wide Reading (WR) and Repeated Reading (RR) will be the focus of this review with an eye on training effects on the execution of higher order skills in monitoring the coherence of a developing mental model of narrative texts.

CHAPTER THREE

FLUENCY AND COMPREHENSION

Introduction

Although the functional relationship between fluency and comprehension is far from clear, there is abundant evidence for a strong association between fluency and reading comprehension across grade levels, ability levels and the comprehension tasks (i.e., cloze tests, standardized reading tests, and free recall) used (Allinder, Dunse, Brunken, & Obermiller-Krolikowski, 2001; Biemiller, 1977-1978; Curtis, 1980; Deno, Mirkin, & Chiang, 1982; Fuchs, Fuchs, & Maxwell, 1988; Jenkins & Jewell, 1993; Therrien, 2004b; Therrien & Kubina, 2006). Coefficients as high as .85 and above have been reported between how many words one reads aloud correctly in one minute and how well s/he does on standardized reading comprehension tests (Allinder, Fuchs, & Fuchs, 1998; Marston, 1989). Albeit very sizeable, the correlational findings are constrained by issues of lack of directionality between the variables in question because bivariate correlational analyses deplete arguments of causality. In the absence of causality, the strength of the relation may be as equally attributable to the effects of reading rate on comprehension as to effects of comprehension on fluency (see Slocum, Street, & Gilberts, 1995, for a review)

In the following sections, I will discuss the underlying reasons for this strong relationship in light of relevant literature and review the extant knowledge base on fluency instruction. This discussion will lead to a focused analysis of comprehension monitoring, a specific area of reading comprehension this study was designed to impact via fluency instruction.

Fluency and Fluency Instruction

Fluency

Research on reading fluency dates back to Cattell's (1886) tachistoscopic work on the unit of recognition and Huey's (1908/1968) emphasis on practice and experience in the unitization of sublexical processes. On the turn of the century, Cattell found that short, familiar words were recognized as equally fast as single letters and that context reduces the time of recognition to as little as 250 milliseconds. Huey, on this note, urged the need to achieve automatic word recognition in order to allocate attention to the construction of meaning. Insights from these pioneers paved the way for research on fluency that was to come only in the cognitive era, after a hiatus of several decades, when cognitive psychology started to regain prominence against behaviorism.

The construct of fluency began to take shape in the late 1960s within the framework of the theories of automaticity (Laberge & Samuels, 1974) and information processing (Norman, 1968; Posner, Lewis, & Conrad, 1972). Cutting across these theories is the notion of automaticity. That is, human information processing requires selective attention for unlearned tasks, which become automatic over trials of consistent stimulus-response correspondence. Once this mapping is learned to the extent to which the task can be executed without controlled attention, its activation from long-term memory is effortless and conscious free (Hasher & Zacks, 1979; Laberge & Samuels, 1974; Logan, 1978, 1997; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977).

Psycholinguistic research is cited to illustrate the autonomy in automatic information processing (Logan, 1997). For example, the Stroop effect has been interpreted as evidence for the autonomy of automatic word recognition. Subjects instructed to name the ink color of written color names are slower to “ignore” the written word and name the color. Stroop interference can also be observed for naming objects with words written across them (MacLeod, 1991). Substantiating the relationship between practice and automaticity is the finding that the Stroop effect displays a developmental pattern through practice. While second- and third-graders do show Stroop effect, the first graders are not susceptible to it (Schiller, 1966). Similarly, L2 learners show Stroop interference in their more automatic first language, but not in a second language they are learning (Tzelgov, Henik, & Leiser, 1990).

According to automaticity theories, a behavior becomes more automatic as a result of practice in consistent environments (Logan, 1979; Shiffrin & Schneider, 1977). While it is believed that practice is a fundamental prerequisite to acquiring automaticity (National Institute of Child Health and Human Development, 2000a), there are differences as to how automaticity is gained through practice. In Logan’s instance theory, for instance, automatization results from multiple recurrences of the instance in which the stimulus was first perceived (1988). The first memory retrieval depends on an algorithmic process (thinking or reasoning) but as the same instance recurs, algorithmic retrieval gives way to automatic memory retrieval which has accrued familiarity with the task and built a reliable and large memory base. However, as few as one instance may suffice to lead to a mental encoding of the input as has been found in studies of Wide Reading which have found equivalent gains in reading rate and greater gains in comprehension compared to simple repeated readings (Kuhn et al., 2006).

In theories of LaBerge and Samuels (1974) and Shiffrin and Schneider (1977), strength of stimulus-response connections are a more important condition to achieving automaticity (J. R. Anderson, 1982; Cohen, Dunbar, & McClelland, 1990; MacKay, 1982) than accumulating instances. Chunking is used to describe how automaticity is acquired in other theories (J. R. Anderson, 1982; Rosenbloom, 1984). In chunking theories, steps to perform a task are reduced to fewer steps or stimulus and response elements are chunked to produce one processing unit for the stimulus response set. As a result, the performance is faster and less effortful.

In reading, automaticity relates to the facility beginning readers acquire in reading single words, usually following a developmental path. To develop the alphabeticity principle—sound-to-symbol rules—of English, students first must be able to manipulate the sounds of the language and build an awareness of their representation by the letters and letter strings (Liberman, Shankweiler, Fischer, & Carter, 1974). This requires children to match 44 sounds to 26 letters and various letter strings. The match between phonemes and graphemes are not perfect; there are exceptions that violate the letter-to-sound rules; words with irregular pronunciations (e.g., pint) and words with equivocal sequences for which more than one sound may apply (e.g., /ea/ in beak, steak, area).

According to Juel and colleagues, without adequate phonemic awareness, acquisition of cipher knowledge (regular letter to sound correspondence) is difficult to develop (Juel et al., 1986). In their model of literacy acquisition, cipher knowledge depends substantially on phonemic awareness. Print exposure is a secondary variable that aids growth in cipher knowledge but it is ineffective without prerequisite phonemic awareness. These predictions were supported by data from a longitudinal study of literacy acquisition by Juel and colleagues (1986). A group of first-graders who had substantial exposure to phonics books but were low on phonemic awareness read only an average of 3.7 nonwords while the high phonemic awareness group with the same amount of print exposure read an average of 27.9 nonwords. Having been exposed to print and phonics instruction for a year in the first grade did not compensate for low levels of phonemic awareness. Exposure to print, however, is expected to contribute to cipher knowledge following a prerequisite degree of phonemic awareness is attained. Theoretically, print exposure provides practice for readers to confirm and consolidate their knowledge of regular grapheme-phoneme correspondence rules (Ehri & McCormick, 1998).

Print exposure makes greater contributions to lexical knowledge, a second component of word recognition skills, which refers to exception words. In fact, print exposure is the only source of gains in lexical knowledge in Juel and colleagues' model of literacy acquisition (Juel et al., 1986). The model assumes therefore that readers learn to read irregular and exception words only through exposure to print. In their longitudinal study of first to second grade, while cipher knowledge was a stronger predictor in the first grade (.208) after the variance it shares with lexical knowledge was controlled for, it explained only marginal variance in second-grade word recognition (.042). A similar but reverse pattern was observed for lexical knowledge after the variance it shares with cipher knowledge was taken into account: lexical knowledge accounted for 3.4 percent of variance in first-grade word recognition, but it explained 20.3 percent of the variance in second graders' word recognition. These findings suggest that second graders, with greater phonemic awareness and cipher knowledge, are more prone to gains from print that aid growth of recognition for exception words.

Stages of fluency acquisition.

The stage models of literacy acquisition provide a more precise picture of how automaticity is acquired in reading words. By the time a normal child starts official schooling, s/he attains a sound grasp of the spoken language system. During the time s/he is taught to read (usually the first grade), the child relies on this knowledge of the spoken language in order to crack the code of reading from print to sounds. S/he faces the task of converting the print into speech by learning correspondences between speech units and written units. From this perspective, learning to read is marked by cracking the alphabetic code for the child. With most of the codes unraveled, the child reads to learn. As s/he keeps reading, use of phonological rules is relegated to less priority. S/he starts building a store of sight words, which could be read instantaneously, obviating any recourse to print-sound rules.

There is wide consensus that phonological system mediates learning to read as reflected in many word reading models. This initial—and usually slow—process, however, gains momentum as the reader reads to learn. The more s/he reads, the more familiarity accrues with sublexical and lexical units (Cunningham, Perry, & Stanovich, 2001; Cunningham & Stanovich, 1991, 1997, 1998). Skilled reading, which an avid reader achieves by doing so, employs mostly direct access of familiar words and word parts in the lexicon. Thus, the phonological mapping of graphemes to phonemes is reserved for unfamiliar words or non-words—words that are not stored in the lexicon.

The brief depiction of a developing reader above is the developmental perspective implicated in print exposure and explained by stages of reading; advanced reading skills develop through successive stages. These stages represent in general the state of reading ability, which hinges on the experiential variables that reside in each stage. Each stage is therefore marked by typical behavior exhibited by most readers.

Developmental theories of word reading recognize that at one point the reader starts building a store of sight words (Ehri, 2005). The start of a sight word collection is assumed by most to follow a maturation level in phonological processing skills. This assumption holds that a reader needs to be proficient at phonological skills before s/he can start to build a sight vocabulary (Ehri, 1995). In a stage model by Frith (1985) there are three stages to fluent word recognition: (a) logographic, (b) alphabetic, and (c) orthographic. In the logographic stage, words are identified based on salient letter and word patterns. In this stage, there is no analysis involved. In the alphabetic stage, letter identity and order are learned by the child. With the alphabetic knowledge increasing, the reader is now able to analyze a word. In the last stage, the analytic knowledge is utilized to a greater degree than the previous stages; orthographic representations are forged for common words or spelling patterns.

Similarly, Ehri's (1995) model of sight words is geared towards explaining the construction of connections between spelling and sounds. The reader starts a stock of "sight" words or word parts as a result of multiple and frequent encounters. Sight reading actually starts in the very first stage of pre-alphabetic stage, parallel to Frith's logographic stage and Chall's (1983) pre-reading stage, as a child makes connections between salient visual features and meaning (e.g., the M arch of McDonalds). The child learns to employ the mediating facilitation of phonology as s/he starts learning about the spelling-sound rules of English through the pre-, partial- and full-alphabetic stages. In the last stage, consolidated alphabetic, s/he is able to use analogies to new words, use the context, and retrieve sight words. Kuhn and Stahl (2000) define sight words as "all those words that have been recognized accurately on several occasions" (p. 414). According to this definition, accurate recognition and multiple encounters with words are presupposed. Readers start a mental representation of words which is first based on their orthography. At each encounter with the word, readers increase their familiarity with the word as they learn more about the word's pronunciation, meaning and spelling through experience.

Prosody.

The discussion of fluency in reading as a progression from accurate to automatic word recognition did not sit well with some for whom there is a further element to consider that is apparently supported by the gains made during fluency training (as in Repeated Readings). Schreiber (1980) argues that automatic word recognition theories, such as the one by Laberge and Samuels (1974), address transition from accurate to automatic word recognition but fail to explain how students are cued to the prosody (syntactic structure of sentences) in print.

According to Schreiber (1980), since print does not signal the prosodic features of speech, students may be at risk for not developing the ability to group words properly and read with expression—as they normally do in speech—with excessive focus on word reading accuracy alone. He, therefore, argues that "it is the ability to compensate for the absence of prosodic cues that enables a reader to achieve reading fluency… by identifying syntactic phrases" that are unmarked in print, unlike phrases in speech which are easy to identify through prosodic cues such as intonation, pitch, and rhythm (p. 182). Schreiber suggests that Repeated Readings leads the child to use other kinds of signals that are specific to print to make up for the lack of prosodic cues. However, Schreiber cautions that it is still not known how the improvements reported by studies using Repeated Readings could be explained by having students read a text repeatedly other than assuming that students will discover "parsing strategies" that they can equate to those used in speech (p. 183).

Schreiber further suggests that students' job of discovering the parsing strategies may be greatly aided if they hear a competent model provide a fluent reading of the passage as is typically done in the listening-while-reading procedure initiated by Chomsky (1978). Subsequent research has confirmed this suggestion by showing significant effects of providing a model rendering of the passage to the students by adults or tape-recordings (Ardoin, McCall, & Klubnik, 2007; Dowhower, 1987; Kuhn & Stahl, 2003; National Institute of Child Health and Human Development, 2000a).

Schreiber's prosody theory is supported by research evidence showing that while accuracy is necessary for the acquisition of fluency, it certainly is not sufficient. Several studies found greater gains in reading rate and comprehension in conditions where students repeatedly read a passage or used context to derive the meanings of unfamiliar words compared to conditions in which they studied lists of words (Dahl, 1979; Fleisher, Jenkins, & Pany, 1979). In Dahl's study, for example, significant gains were observed only in the conditions where struggling second-graders read running text for either repetition or context clues. However, there is also evidence that speed of single word reading is strongly related to reading ability and comprehension (Biemiller, 1977-1978; Perfetti & Hogaboam, 1975), which may be a direct reflection of reading ability of good readers, not a reflection of training of reading single words.

In a more direct assessment of practice effects on expressive reading that Schreiber attributes to Repeated Readings, Herman (1985) described gains in accuracy and rate and a decrease in miscues at posttest. Despite these gains, a decrease in pausal intrusions were only noted within practiced passages; students continued to show a constant occurrence of pauses that interrupted idea units on new passages—between-passage declines were however later reported by Dowhower (1987). Although decrease in pauses on practiced passages failed to carry over to unpracticed passages, Herman suggests a decrease in the length of pauses as one plausible reason for increases in rate and increase in an indirect assessment of comprehension which included appropriate semantic and syntactic miscues as well as the correctly read words; however, she does not provide data for this claim. Shorter pauses may indicate that less time was needed by readers to process the text.

Dowhower's (1987) study of prosodic improvement in transitional second-grade readers who participated in assisted and unassisted Repeated Readings documented the salutary effects of the practice provided. The transitional students (slow but accurate decoders) exhibited gains in expressive reading in addition to gains in rate, accuracy, and comprehension; both groups (assisted and unassisted) read with longer phrases and fewer pausal intrusions. In addition, all students exhibited an increased number of sentence-final word lengthening and used falling pitch for these words, appropriate prosodic features. Despite these similar gains, it was the assisted group who read with fewer pausal intrusions than the unassisted group. Based on these findings, Dowhower concludes therefore that “as Schreiber [1980] suggested, Repeated Readings helped children tacitly develop prosodic strategies for organizing text. Even though the words were written one-by-one on the page, the students began perceptually to isolate phrases with intonation, segmental lengthening, and appropriate pauses” (p. 403). That is, the students learned to approximate the tone of the author as they started to use appropriate phrasing and intonation.

Recent studies identified specific components of prosody in relation to decoding and reading comprehension in second- to third-graders. Schwanenflugel and colleagues (Miller & Schwanenflugel, 2006; Schwanenflugel, Hamilton, Kuhn, Wisenbaker, & Stahl, 2004) found that measures of pitch changes explained unique variance in reading comprehension independent of decoding. Appropriate pitch changes accounted for 6.7% of residual variance after common variance with decoding was removed. Unlike appropriate pitch changes, measures of appropriate pausing failed to explain unique variance in reading comprehension over and beyond that explained by decoding. Based on these findings, Schwanenflugel and colleagues conclude that pauses are indicators of decoding problems and appropriate pitch changes are predictors of reading comprehension.

These findings converge on a premise that expressive reading may mature slowly, through practice and experience, a conclusion that was drawn by Schwanenflugel, Hamilton, Kuhn, Wisenbaker, and Stahl (2006). In this study data on word level fluency, automaticity (as operationalized by Stroop task), and text level reading fluency (expressive reading) supported only a *simple reading fluency* model as opposed to the *text reading fluency as mediator* model. The *text reading fluency as mediator* model represents phrase-level fluency skills and is assumed to account for a larger portion of reading comprehension than single word reading (Schreiber, 1980, 1991). The data was better represented by a simpler version of fluency which comprised both word and text fluency within the same construct. Thus, it appears from this study that for elementary school students a mediating fluency factor between word and reading comprehension may require acquisition of greater proficiency in single word reading. Parallel results are reported from an earlier German study by Kowal and colleagues (Kowal, O'Connell, O'Brien, & Bryant, 1975) in which second graders were observed to be processing text word by word while fourth graders were reading in phrases despite several pauses.

In summary, early definitions characterized fluency as accurate and fast word recognition (Laberge & Samuels, 1974; Perfetti, 1985) influenced by the assumptions of the cognitive theories of information processing. In earlier models of reading, decoding and comprehension are dissociable and interdependent (Gough et al., 1996). Neither is sufficient alone for reading to occur. This simple view of reading and automaticity models saw comprehension as attention-demanding and word recognition amenable to becoming automatic with practice. However, the effects of practice were limited to single word reading. Practice was confined to word recognition only, and its effect did not extend to appropriate phrasing and use of prosodic features commensurate with speech. Recent conceptualizations view fluency as a bridge between decoding and reading comprehension (Dowhower, 1987; Pikulski & Chard, 2005; Schreiber, 1980; Wolf &

Katzir-Cohen, 2001) and emphasize the "simultaneity of decoding and comprehension that is the essential characteristic of reading fluency" (Samuels, 2007, p. 564).

When decoding and comprehension are not bridged by fluency, comprehension difficulties persist in the absence of decoding problems. In other words, poor comprehension is accompanied by disfluency in spite of grade level decoding. This was illustrated by a group of ninth-grade poor comprehenders in an urban school who demonstrated grade-level decoding skills but were two years behind in fluency (Rasinski et al., 2005). A close relationship between fluency and reading comprehension was also the finding of a NAEP study by Pinnell and colleagues (1995), who found that about 44% of the fourth graders were disfluent with grade level text and failed to comprehend texts at their grade.

Fluency Instruction

Among the first to put automaticity theory into a practical application are Dahl (1974), Samuels (1979), and Chomsky (1978), who developed instructional methods that induce practice and lead to automatic word recognition and faster reading. Dahl and Samuels' repeated reading and Chomsky's reading-while-listening methods are to date the first examples of fluency training ready for intervention.

In his now classic study of 1974, Dahl compared the effectiveness of repeated readings on four poor second graders who took part in the fluency training to another group of poor second graders who did not over the course of eight months. At the end of the training, students' reading rate, comprehension ability, and miscues were collected. The reading rate was significantly increased for the treatment group who also read with fewer miscues than the control students. On a standardized reading comprehension test, comprehension did not distinguish the groups. However, treatment effects were found on a cloze measure.

In her listening while reading version of repeated readings, Chomsky (1978) had five third-grade students listen to and repeatedly read stories until they read the stories without any errors. The stories were played over a tape player, which the students used individually. At the end of 15 weeks of “mimic” training, students gained up to six months in word recognition and up to a year in oral reading speed. In fact, it was later found that repeated readings alone and reading-while-listening contribute equally to improvements in fluency; Rasinski (1989) found both methods effective in improving third graders’ accurate and fast reading.

Repeated Readings “is a supplemental reading program that consists of rereading a short and meaningful passage until a satisfactory level of fluency is reached” (Samuels, 1979, p. 404). Recent reviews confirmed that the program is effective for both nondisabled and disabled readers, with larger effects for fluency than comprehension (Meyer & Felton, 1999; National Institute of Child Health and Human Development, 2000a; Therrien, 2004a). The Repeated Readings program was employed to deliver fluency instruction to one of the treatment groups.

In its initial implementation, Repeated Readings involved students working on the passage until the preset criterion rate was reached (Dahl, 1979). “To move to the next passage, it was not necessary to have recognized all the words with 100% accuracy but it was necessary to reach the speed criterion” (Dahl, 1974, p. 8). Therefore, in initial definitions of fluency stronger emphasis was placed on speed of reading with adequate comprehension accuracy.

In addition to these initial versions of repeated readings, other means were also used to engage poor readers in focused practice, such as having students read passages in which phrases are demarked (Carbo, 1978) or having students follow along a song while reading the lyrics (Newsom, 1979). Despite inadequacies in data reported by some of these researchers, it appears that gains in word recognition is a common outcome in fluency programs that use repeated readings as a means of practice. Comprehension, however, does not seem to be affected as much, a contention that the National Reading Panel approved in a review of fluency studies with a set of stringent meta-analysis criteria. (The NRP results will be reviewed in greater detail below.)

Several components of fluency instruction have also been credited for the effectiveness of the program. For example, researchers have reported consistent improvements following interventions that include a model (Chomsky, 1978; Conte & Humphreys, 1989; Strong, Wehby, Falk, & Lane, 2004). The models can come in the form of an adult or a partner reading the passage to the struggling reader; or an audiotape of good reading could be used as a model prior to the practice of repeated readings (Chomsky, 1978; Hollingsworth, 1978; Young, Bowers, & MacKinnon, 1996). Providing corrective feedback also seems to increase the effects of a fluency program (Mercer, Campbell, Miller, Mercer, & Lane, 2000; National Institute of Child Health and Human Development, 2000a; Samuels, 1979; Strong et al., 2004). Setting the number of rereads at four seems to be sufficient to make gains even though the first implementation of the program by Samuels (1979) enforced a set criterion—reaching a target rate of words read correctly per minute (Meyer & Felton, 1999; Therrien, 2004b).

Reading practice can also be provided with less intervention in settings wherein students read connected text independently. During a specified period of time, students engage in “sustained silent reading” of relatively easy material that they choose. Reading material at their independent level, students “overlearn” linguistic structures they already know, focus on comprehension, learn to adjust their reading rate contingent upon the purpose they set, expand their world knowledge, build self confidence that leads to greater desire and interest in reading (Durkin, 1993; A. J. Harris & Sipay, 1990). Because enjoyment is the goal of sustained silent reading, students are not burdened with challenging text that may slow down their reading. In independent reading, students’ attention shifts from deciphering (unfamiliar) words to comprehension. Students are not forced to focus their attention on hard vocabulary and content but enjoy a text at their comfort level.

Independent reading is considered an important factor for developing important reading skills while providing practice and enjoyment to the reader. An early start at decoding has been shown to be related to increased reading exposure in the following years (Cunningham & Stanovich, 1997). Cunningham and Stanovich showed that students with high scores on first-grade decoding, word recognition, and comprehension measures became avid readers in eleventh grade. Although the impact of print exposure on fluency was not directly addressed by Stanovich and his colleagues, most agree that the relationship between fluency and print exposure is reciprocal as well (Stecker, Roser, & Martinez, 1998); that is, while fluency could be conceived of as the antecedent to avid reading, it could also be a reflection of reading widely. Although not direct, there is evidence that print exposure has unique influences on spelling skills independent of third variables that it shares variance with, such as I.Q. and comprehension (Cunningham & Stanovich, 1991; Cunningham & Stanovich, 1993; Cunningham & Stanovich, 1998). When students read independently, they come across frequent language structures (including letter patterns, words, and phrases) and increase their automaticity of reading these structures. Several studies have shown positive effects of independent reading on reading rate (Dwyer & West, 1989, 1994; Mathes & Fuchs, 1993) and comprehension (Kuhn, 2005b).

Independent reading has also been attributed with positive gains in vocabulary knowledge (Nagy, Anderson, & Herman, 1987), in reading comprehension (Reutzel & Hollingsworth, 1991), and reading rate (Dwyer & West, 1989, 1994; Mathes & Fuchs, 1993). In a rate enhancement program implemented by Dwyer and West (1989), college readers completed a six-week reading program at a rate of 348 words per minute, an increase of 138 wpm, while comprehending most of the material (78%). Although lower, the second Dwyer and West study (1994) reported a final reading rate of 278.3 wpm at the end of 25 days of 15-min of Sustained Silent Reading by 76 college students. Compared to the starting rate of 242.15 wpm, the final gains are substantial and statistically significant ($F = 16.38$, $p < .05$). Further analyses revealed a linear trend of increasing rate from the first to the fifth week, as well as a quadratic trend due to smaller gains after the second week, consistent with Logan's (1997) argument that increase in rates levels off after a certain threshold is reached. Dwyer and West conclude, therefore, that "evidence reported herein suggests that providing time and otherwise encouraging normal reading (SSR) promotes reading rate" (1994, p. 11).

The efficacy of these fluency procedures was evaluated in a painstaking review undertaken by the National Reading Panel committee of fluency training studies from both education and psychology journals. A selection of 16 (guided) oral repeated readings studies which met the stringent criteria of including a treatment and a control group with pre- and post-test data was subjected to a meta-analysis. A mean effect size of 0.41 was obtained from 99 effect sizes calculated by comparing treatment groups' gains to control groups' in fluency programs providing practice through repetition. While the effects were higher for measures of word knowledge (.55), and fluency (.44), the mean weighted effect size for comprehension was lower, if not substantially, .35. A more focused meta-analysis was conducted by Therrien (2004) on the effects obtained from studies using pure repeated readings. Therrien's review of repeated readings studies reported an average effect size of .95 for fluency gains and .71 for comprehension gains on the non-transfer passages—passages that were read repeatedly. In transfer passages, the mean effect sizes were lower: an ES of .50 was found for fluency gains and .25 for comprehension. While the effects are greater for fluency, they are in the moderate range for comprehension.

As for independent reading, the NRP report was reluctant to indicate a position on its effectiveness, except for calling for more research. The report did not find any effects of encouraging students to read independently; neither of the programs reviewed, i.e., Accelerated Reader, Sustained Silent Reading, showed improved comprehension relative to a control group. Although independent reading failed to garner the support of the NRP report, many in the field have encouraged teachers to provide children with ample time and opportunities for independent reading. Adams empathically argued that "beyond the basics, children's reading facility, as well as their vocabulary, and conceptual growth, depends strongly on the amount of text they read" (1990, p. 127). Cognitive consequences of reading exposure are elaborately demonstrated in the work of Stanovich and his research team, who have proposed that avid readers are engaged in a positive feedback loop with concomitant effects for improving basic reading skills and leading to greater print exposure (Cunningham & Stanovich, 1991, 1997). Independent reading has therefore been recommended for students with adequate decoding skills who are more likely to make larger gains in vocabulary and comprehension than less skilled decoders for whom more practice in word recognition (through a repeated readings program) may be more beneficial (Pikulski & Chard, 2005).

Unlike the NRP findings, several studies and reviews have challenged repetition as an essential means of practice for disfluent readers. Kuhn and Stahl (2003), reviewing the literature, noted that non-repetitive, wide reading may lead to gains commensurate to those from repeated readings and that reading rate gains from repeated reading may not generalize to unpracticed texts. In addition, some research has shown that older, struggling readers may benefit equally from the same amount of wide reading (non-repetitive) to improve on rate, word recognition, and comprehension (Homan, Klesius, & Hite, 1993; Rashotte & Torgesen, 1985). Furthermore, a study by Mathes and Fuchs (1993) revealed larger gains in reading fluency from a sustained reading condition than a repeated reading condition in intermediate grade students with reading difficulties. There were no differences between the groups on comprehension.

Recently, a comparative study by Kuhn (2005) implemented repeated readings and a non-repetitive wide reading intervention program with second graders. Both

Repeated Reading and Wide Reading students improved in rate (including isolated word recognition measured by TOWRE and prosody measured by NAEP's 1995 Oral Reading Fluency Scale), with the Repeated Reading students having an edge over the WR students. Students in the WR program, however, were better at comprehension than the Repeated Readings group. Although Kuhn argues that students might have developed a program bias for the intervention program they were assigned to (RR students were cued to read for fluency; WR students for comprehension, see O'Shea, Sindelar, & O'Shea, 1985 for the differential effects of cuing for rate vs. cuing for comprehension), Kuhn's findings are consistent with a later study by Kuhn and colleagues (2006), who reported that higher comprehension gains for the Wide Reading students might have stemmed from exposure to a greater amount of text. Students in this program read a total of 18 passages compared to eight passages read by the Repeated Readings group.

Kuhn and associates (2006) implemented Fluency Oriented Reading Instruction (FORI) and Wide Reading to a large sample of second grade students from two major sites in the U.S. The FORI instruction involved a weeklong lesson plan which included teacher modeling, repeated reading, partner reading, choral and echo reading as well as comprehension activities on one passage. Students in the Wide Reading classes read three different passages each week. Like students in the FORI classes, Wide Reading students were provided scaffolding.

The results found no significant differences between the two programs on word reading, oral reading fluency and reading comprehension; however, compared to the control groups, Wide Reading led to better and quicker gains. Students in the Wide Reading classes showed the significant gains (relative to the control students) in word reading efficiency and reading comprehension as soon as the mid-year assessment point. The FORI students caught up with the Wide Reading students at the end of the year. The FORI students never reached the Wide readers in oral reading fluency; they never showed larger gains than the control group. The authors refer to the *instance theory* (Logan, 1988) to explain the Wide Reading group's gains in oral reading fluency: students in this group accrued a greater number of traces at the lexical, phrasal, and textual levels than the FORI students who were exposed to only one third of the texts that the Wide Reading group read. Similar gains across the two groups in single word reading were assumed to reflect the generality of sublexical traces across passages at this level. It appears that one reading is sufficient to encode sublexical units per week (the Wide Reading group), as is suggested by the instance theory (Logan, 1988).

A text difficulty explanation was also offered for the larger gains displayed by the Wide Reading group. Although texts used in the fluency programs were above grade level for most students and appropriate scaffolds were provided, the wide reading group made greater gains in reading rate and reading comprehension. Challenging texts and a wide variety and breadth of words, concepts, and syntactic structures might have contributed to increases in the reading comprehension, and oral reading fluency of the Wide Reading group. This "wide" exposure consequently showed positive effects in oral reading fluency, the acquisition of which eluded the FORI group.

Prevailing in the literature reviewed here is the conclusion that gains produced by practice-inducing training programs seem to be limited to increased skills in word recognition. The fact that equally meaningful gains did not translate to comprehension encourages reading researchers to break down the construct of comprehension and examine components of comprehension with fine-grained methods. The next section reviews the literature on constructing coherent mental models and the paradigm that has been used to tap readers' awareness of their online comprehension.

Comprehension Monitoring

In the resurgent cognitive era, a new component of metacognitive processes was added to our understanding of reading comprehension: comprehension monitoring (Sternberg & Powell, 1983). It is considered the evaluation step, which presumably precedes a more active step of regulation in which the comprehender plans and implements a set of actions to fix a comprehension break (Baker, 1985, 1989). That is, before the reader can take corrective action, she needs to first be aware that her comprehension is not going right. A break in comprehension may flag her even though she may not know what caused the break.

Baker (1985) identifies three main sources of comprehension problems that may flag the reader; for each the reader employs a specific standard: lexical, syntactic, and semantic. A difficult word or a non-word may cause the difficulty in comprehension, for which the reader uses a lexical standard. Using this standard the reader evaluates her understanding of the individual word independent of the context. The reader may be prompted to use a syntactic standard for a set of scrambled phrases (Paris & Myers, 1981). She recognizes that the sentences contain scrambled phrases using her grammatical knowledge, the syntactic standard. Comprehension problems may also be caused by inconsistencies in the text, for which the reader invokes semantic standards. Of several semantic standards that Baker identifies (e.g., propositional and textual cohesiveness, external consistency, internal consistency, and informational clarity), internal and external inconsistency standards have received more interest in the research literature. According the Baker (1985), the reader uses the external consistency when she notices that ideas in the text violate a known fact and she uses an internal standard to check whether ideas are consistent with one another.

To examine the developmental pattern of the standards and skill differences in their use, researchers have created the error detection task by inserting inconsistencies in passages. In an externally inconsistent passage, a proposition asserted in the text contradicts the reader's general world knowledge (e.g., most dogs meow, Ruffman, 1996, p. 35). In a logically inconsistent passage, a sentence contradicts another sentence (e.g., most people I know like corn, most people I know do not like corn, Ruffman, 1996, p. 35). Inconsistency could be caused by a long, difficult word, which may undermine a child's ability to construct a "definite conclusion about what the text is about," for example (Ruffman, 1996, p. 35). Of the three, comprehension monitoring may be better measured in texts with logical inconsistencies or lexical ambiguities than external consistencies; detection of external consistencies may lead to problem identification instead of comprehension monitoring if students are asked to find out what is wrong with the text (Ruffman, 1996). While most readers can detect the lexical inconsistencies, logical and external inconsistencies are easier for older and good readers (Baker, 1985).

Subjects' verbal responses or reading times of sentences may be collected to tap their ability to detect errors. After reading the text, students may indicate verbally whether or not the story makes sense (August, Flavell, & Clift, 1984; Baker, 1979; Baker & Brown, 1984; P. L. Harris, Kmithof, Terwogt, & Visser, 1981; Markman, 1979), or read passages one line at a time on the screen of a computer, which stores reading times on each line (Baker & Anderson, 1982). Data collected through verbal protocols are known to have a number of flaws, due to which the task may not reveal online cognitive processes: students (a) may fail to report the contradictions because they may try to "repair" it by making inferences to resolve the error (Baker, 1979; Baker & Anderson, 1982; Winograd & Johnston, 1982), (b) may choose not to "question" the author as in the case of most young readers (Robinson & Robinson, 1977a), or (c) may fit the developing mental model around the contradiction, and thus misrepresent it (Baker, 1985; Vosniadou et al., 1988; Winograd & Johnston, 1982). In studies using verbal protocols, therefore, students are told beforehand that there is something wrong with the way the story is written in order to encourage students to actively monitor their comprehension (Markman, 1979).

A study by Harris and colleagues (1981) perhaps best illustrates the inadequacy of oral protocols, especially with young children. The researchers gave one group of children texts in which the title was inconsistent with a line in the text. A second group was given the same text in the consistent condition; the title was consistent with the rest of the passage. Data from verbal protocols and reading time did show a different pattern of results. While 11-year-olds were significantly better at recognizing the inconsistency in verbal protocols than the 8-year-olds, both groups were equally slower reading the line that was inconsistent with the title. These findings imply that the younger group was engaged in constructing a coherent mental model while they were not explicit about the process.

Supportive evidence for the use of reading time data is also reported by Baker (1979) and Baker and Anderson (1982). In Baker's study, college students were instructed to read expository passages and answer discussion questions about the passages. The passages were inconsistent due to an inappropriate logical connective, an ambiguous referent, or inconsistent information. The recall data indicated that substantially large amount of text confusions were not detected by the subjects (62%). It was discovered, however, in the protocols that students tended to resolve the inconsistencies when they encountered the inconsistencies by using fixup strategies. For example, they made inferences to repair the break caused by the inconsistent information, used their prior knowledge to fill the gap, and reread the passages or looked ahead for clarifying information. Or, some students decided not to dwell on the comprehension break and read along. In other cases, they thought that they had understood the passage by erroneously calibrating their comprehension.

In the Baker and Anderson (1982) study, college students read passages on a computer screen presented one sentence at a time. Students read the passages at their own pace and were allowed to look back at previous sentences if they needed to. After reading the passages, students were asked to indicate which sentences were inconsistent. Half the students knew before the experiment about the existence of inconsistencies in the passages while the other half did not. The reading time data revealed that students were slower to read the inconsistent sentences, and they regressed the earlier portions of the inconsistent passages more often than the consistent passages. The behavioral data were paralleled by the retrospective reports in which students detected a higher proportion of the confusions than the students in the Baker (1979) study. Moreover, there was no difference between the groups of students who were told in advance about the existence of inconsistencies in the passages and those who were not. It appears from these results therefore that there is a tendency among adult readers to monitor their comprehension naturally in the absence of specific instructions (Baker & Brown, 1984).

It has been found in error detection studies that there are age and skill differences in successfully evaluating one's state of comprehension. Usually younger and poor readers are less sensitive to logical contradictions than older and good readers (August et al., 1984; Baker, 1979, 1985; Baker & Anderson, 1982; Garner & Reis, 1981; Markman, 1977). Children tend to detect contradictions involving words or a known fact more often than logical contradictions. The processing requirements involved in detecting logical vs. external inconsistencies may account for this difference: subjects (both children and adults) are usually fairly good at checking an externally inconsistent proposition to their world knowledge, if they possess the requisite world knowledge. However, they have to rely on a weaker standard in logical inconsistencies: they have to match the inconsistent propositions to one another (Baker, 1984a) to detect the error in the passage. Another explanation for the difficulty with logical inconsistencies is working memory limitations and processing efficiency. Children or less skilled readers may not possess sufficient working memory resources to represent and compare the inconsistent propositions simultaneously (Yuill & Oakhill, 1991; Zabrucky & Moore, 1989), the differences may be greater when the distance between the logically inconsistent propositions are spread out (Ackerman, 1984; Oakhill et al., 2005).

The logical error detection task taps the ability to represent propositions containing the inconsistency and to compare them during reading (Markman, 1979). Vosniadou and colleagues found a strong relationship between elementary school students' recall ability and detection of (internal) inconsistencies in short stories; that is, when students recalled the inconsistent propositions, they were more than likely to detect the inconsistency (Vosniadou et al., 1988). In other words, detection depended on the extent to which the inconsistent information was represented. Also, the detection rate was higher in the listening condition compared to the reading condition. For students in early grades (i.e., first vs. third vs. fifth) detection was easier when the stories were read to them; however, the difference between reading and listening seemed to disappear with increasing grade. This finding implicates the processing efficiency that accrues at higher grades in reading, although no information was collected on students' working memory capacity except for overall reading ability. Unlike Vosniadou et al.'s (1988) findings from elementary school students' verbal protocols, recent reading time studies have yielded data indicating that poor readers may be able to reactivate relevant information but fail to integrate it with the target sentence that is currently in working memory, thus fail to exhibit a tendency to resolve the inconsistency (Albrecht & O'Brien, 1993; Cook et al., 1998; Long & Chong, 2001).

Manipulation of the distance between inconsistent propositions in the text has also proven to differentiate skilled from unskilled readers, leading to a distance by skill interaction, which is taken as a clear indication of working memory constraints in less skilled readers. While there is no difference between skilled and unskilled readers in the near condition wherein the inconsistent sentences are adjacent, only skilled readers detect the inconsistencies in a global condition in which the inconsistent sentences are separated out by several intervening sentences (Oakhill et al., 2005). The close and distant conditions have been argued to the tap into different representations of the texts read. In the close condition, the critical information sentences are presented in close proximity and are supposed to remain active in working memory during processing. Conversely, the two critical sentences do not reside in the same processing cycle in the distant condition. Slower reading times in the distant condition are taken as evidence for construction of a mental model, which requires sensitivity to information from earlier parts of the text, which may be reinstated during comprehension.

It appears that less skilled readers do not have difficulty mapping incoming information onto their developing mental model when the memory demands are low (adjacent condition). Thus, less skilled readers can monitor their comprehension over short segments of text but fail to do so when conflicting sentences are separated out by several intervening sentences, due to their limited mental capacity. In this sense, they seem to be able to build only partial and incomplete mental models, unlike the skilled readers whose mental representations seem to be completely integrated (Garnham et al., 1982). For example, in a study by Oakhill and colleagues (Oakhill et al., 2005), 9- to 10-year-old skilled and less skilled students read sentences such as “Gorillas sleep on the ground on a bed of leaves and they like to eat different types of fruit,” “Gorillas sleep in trees and they often build shelter out of leaves above them, to keep the rain out” in short passages and indicated whether or not the sentences made sense after reading them. The results revealed a distance by skill interaction effect: less skilled readers did far worse on the distant condition than skilled readers whereas the two groups did not differ in the adjacent condition.

Although several researchers have argued that readers who fail to detect inconsistencies in text do not actively engage in building coherent mental models (Markman, 1977, 1979), evidence is mounting on an alternative view: nondetectors may not be able to integrate propositional representations that they seem to adequately construct. Among a number of theories regarding the underlying factors for skill differences in error detection, recent studies using reading time paradigms have leaned towards working memory limitations and building incoherent mental models (Cook et al., 1998; Long & Chong, 2001). There is a tacit conjecture that deficient processing skills may undermine students' ability to integrate new information onto a developing mental model and monitor their comprehension more directly.

For example, in Albrecth and O'Brien (1993), one passage introduced a protagonist called Mary who was waiting for a friend at a restaurant. Mary is described either as a junk food addict or a strict vegetarian. Following the introduction of the protagonist's characteristic, several filler sentences were presented in order to remove information about Mary from working memory. The filler sentences were then followed by a target action which described Mary ordering a cheeseburger and fries, which was consistent with her junk-food addict characteristic but was inconsistent with her description as a strict vegetarian. While the target action of ordering a cheeseburger and fries was locally coherent (consistent with the immediately preceding filler sentences), it was globally incoherent; that is, it was inconsistent with the introductory information in the inconsistent condition.

A sample of normal college readers experienced comprehension difficulty while reading the target sentence of the inconsistent condition, indicated by their slower reading times relative to the target sentence in the consistent condition (Albrecht & O'Brien, 1993). When reading the target sentence, concepts such as Mary and ordering junk food activated earlier statements about Mary's reading habits. According to the Resonance

Model, reactivation of earlier information occurs in both consistent and inconsistent conditions; however comprehension difficulty arises only in the inconsistent condition. The inconsistency effect has been robustly replicated by other researchers, such as Cook and colleagues (1998) and Long and Chong (2001).

In a college sample of good and poor comprehenders (based on Verbal SAT scores), Long and Chong (2001) extended Albrecht and O'Brien paradigm by incorporating variables of skill (good vs. poor readers) and distance (global coherence vs. local coherence). The materials were passages used by Cook and colleagues (1998) with the structure described above for the Albrecht and O'Brien (1993) study (an example passage is attached in Appendix A). To create the local condition in their study, Long and Chong condensed the filler section of Cook and associates' passages to one sentence, leaving the global condition's filler section at mean 77.6 words and local condition's at mean 15.1 words. Each passage started with an introduction of two characters (Ken and Mike), continued with elaboration of one of the two characters, with either a short filler (15.1 words) or a long filler (77.6 words), the target action performed always by the first character (Ken), and ended with a brief conclusion. Good comprehenders experienced difficulty at both the local and global conditions for only the relevant character (Ken). Poor comprehenders in contrast exhibited an inconsistency effect only at the local condition for both characters, suggesting that "their text-based representation was inaccurate, or at least incomplete, with respect to their representation of the two characters" (p. 1426). That is, although the poor comprehenders were able to detect the inconsistency in the local condition, they failed to differentiate between the characters.

They should have showed an inconsistency effect only for the appropriate (the first character), not for both characters.

Long and Chong (2001) further pursued the hypothesis that poor comprehenders failed to reactivate character elaboration once it was backgrounded by a lengthy filler section (in the global condition) via a probe-verification task. Probe sentences containing information about the character description were presented either after the character description, before the target action, or after the target action. Both good and poor comprehenders were equally faster to respond to the probe after the description section and the target sentence. Consistent with the Resonance model, both first and second-character elaborations were activated at the target sentence, even though only the first character description was relevant to the action. This finding rejects the hypothesis that poor comprehenders fail to maintain global coherence because they lose access to information earlier in the text. Poor comprehenders had no difficulty reinstating prior text information into memory and "character description was sufficiently well encoded in poor comprehenders' text representations to be reactivated by information in the target action" (Long and Chong, 2001, p. 1428).

Long and Chong (2001) argue that poor comprehenders' difficulty is at the discourse level, where integration is required between propositions even when each proposition may be encoded well enough to be reactivated from long-term memory. In this sense, their findings are consistent with other studies showing that poor comprehenders experience difficulty constructing globally coherent text representations (Cain et al., 1998; Garnham et al., 1982; Long et al., 1994; Long et al., 1997; Oakhill et al., 1986).

Summary

In sum, despite the buzz around "active reading" (Durkin, 1978), the field is amiss on how to measure active reading and how to induce such metacognitive skills in poor readers. In general, comprehension is treated as a single, unitary construct that is measured with standardized or experimenter-developed tests, and the various methods used (e.g., strategy training and fluency-enhancing training programs) are still far from unequivocal empirical support. Furthermore, recent investigations overall report equivalent gains in comprehension from fluency programs, with gains favoring non-repetitive Wide Reading procedures as opposed to Repeated Reading procedures in only one study with second-grade transitional readers (Kuhn, 2005a). Consequently, further research is warranted to identify specific effects on comprehension (i.e., establishing coherent mental models) that result from training.

This study aimed to test the hypothesis that fluency increases improve struggling college readers' ability to integrate sentence meanings in discourse and to build a coherent mental model of the text as a whole.

CHAPTER FOUR

METHODOLOGY

The previous chapters described the theoretical underpinnings of mental models and the consequences of inefficient processes on higher-order comprehension skills such as establishing and maintaining coherence. Also reviewed were approaches to promoting reading fluency in struggling readers through practice. Building on this review, this chapter discusses preliminary data gathered from a pilot study in Spring 2009 and details the design and the procedures of the fluency intervention that was conducted with college readers in Summer 2009.

Pilot Study

A pilot training program was conducted in Spring 2009 to estimate the adequate dosage of fluency instruction with struggling college readers by identifying the session beyond which fluency gains started to level off. This information was used to determine the duration of the fluency intervention program implemented with study participants in Summer 2009.

Participation of college students was solicited at a 4-year university (Site 1) and a 2-year community college (Site 2) via a study brochure (see Appendix B). Fifteen responders were identified as struggling based on their performance on the ND Reading Comprehension subtest. Additional reading measures were also administered to Pilot participants; however, none of the Pilot data was used in any of the analyses conducted in this study.

Struggling readers were enrolled in a Reading Course at Site 1 and a Developmental Reading class at Site 2. The Reading Course is designed to give practice for the Reading Test, a high-stakes reading comprehension test. All undergraduate students in the state are required to pass the Reading Test in order to graduate. The Developmental Reading course, which was offered at Site 2, is a lower level course designed to prepare lower-achieving, developmental readers for the Reading Course so that they could pass the Reading Test. While three participants spoke English as their first language, the rest of the sample spoke Bosnian (1), Creole (2), and Gujarati (2) as their primary language. Table below lists the Pilot participants' characteristics.

Table 2

Pilot Participant Characteristics

Site 1 (n = 8)	
Age	*n* = 8; *M* = 24; *SD* = 4.47
Male	3
Female	5
Site 2 (n = 7)	
Age	*n* = 6; *M* = 21.67; *SD* = 6.59
Male	2
Female	5

The struggling students were randomly assigned to either a RR (*n* = 6), a WR (*n* = 5), or a No-Treatment Control condition (*n* = 4) and followed their respective study conditions individually. Twenty-one sessions of training were planned for the fluency trainees in the Pilot study; however, only 3 students completed the 21 sessions, with the number of sessions completed varying from 2 to 21. A minimum of 16 sessions were completed by 4 WR students and only by 2 students in the RR condition.

Participants engaged in fluency intervention (i.e., Repeated Readings or Wide Reading) individually with the experimenter for about 30 minutes. Although initially the intervention was planned to take place three times a week, not all participants were able to comply with this study provision. This compromised the systematic delivery of fluency intervention and resulted in variation in the weekly dosage of intervention received by the participants. Only three students completed the training three times a week; other students' attendance varied from one to three days of training a week.

A binder was compiled for each participant. Binders for the fluency intervention students were compiled with grade-level passages from the *Timed Readings* (Spargo & Williston, 1975) series. In addition, the binders included (a) the Answer Key, (b) the Progress Graph (see Appendix C), (c) the Time-Rate Conversions Table (see Appendix D), and (d) maze tests. Passages from the next-grade level were included in the binders when the participants met the instructional criteria, which were instituted to advance them to more difficult passages. In order to meet the criteria, participants were required (a) to read at 400 wpm on three consecutive readings and achieve 80% comprehension accuracy per session and (b) to repeat this performance over three sessions in a row. Criterion *a* was recommended by the authors of the *Timed Readings* (Spargo & Williston, 1975) series. Criterion *b* was instituted to ensure adequate exposure to grade-level passages before advancing to next-grade level passages in the series. Two participants met the instructional criteria and continued their training on passages from the next grade level.

Fluency training took place in the library of the experimenter's academic department with Site 1 participants and in a small tutoring room in the Learning and Tutoring Center of the community college with Site 2 students. Repeated Readings participants read a grade-level passage four times back to back, each time recording their reading rate on the Progress Chart. Following the fourth reading, they answered 10 comprehension questions about the passage and recorded their percent correct score on the Progress Chart. Intervention was conducted with participants reading silently, and no feedback was offered on the pronunciation or meaning of unfamiliar words.

Wide Reading participants read four different grade-level passages each once, also silently. When finished reading the passage, they plotted their reading rate on the Progress Chart for that passage, and answered two literal (questions 4 and 5) and one inferential (question 6) comprehension questions. Only three questions were required of the Wide Reading participants per passage. This was done to hold the number of questions answered by the groups as equivalent as possible and thus to prevent the Wide Reading participants from gaining an advantage over the Repeated Readings group in reading comprehension as a result of answering more questions. In all, 10 comprehension questions (5 literal and 5 inferential) were answered by the Repeated Readings group and 12 comprehension questions (8 literal and 4 inferential) were answered by the Wide Reading group per session.

The decision of requiring WR students to answer two literal questions was made to ensure that the Wide Reading students read the passages carefully. Because the passages in the *Timed Readings* series deal with topics of general knowledge, participants may develop a tendency to rely solely on their background knowledge to answer the comprehension questions if they are to answer only or mostly inferential questions. By having participants answer more literal questions than inferential questions, it was expected that the participants would not skip over important information and construct complete mental representations for the passages they read. Cumulatively, however, this strategy resulted in a disproportionate number of literal (8) versus inferential (4) questions answered by the Wide Reading participants; three literal questions more and one inferential question less than the Repeated Readings group.

Linear Trend Analysis was used to estimate adequate dosage of training on the reading efficiency scores of participants. This analysis helped identify the point in training at which gains started a stable linear trend. Reading efficiency scores were derived from the product of reading rate (words per minute) and comprehension accuracy scores (percent correct), following Carver's (1990) definition. Percentage of comprehension accuracy scores were taken in order to yield a metric of comparison between the Repeated Readings and Wide Reading groups; the former answered a total of 10 questions per session while the latter answered 12 questions. Mean reading efficiency scores from six participants (2 Repeated Readings; 4 Wide Reading), who completed at least 16 sessions, were used in the Linear Trend Analysis, which plotted the mean reading efficiency scores against sessions (N = 16).

A visual inspection of the resulting Linear Trend Analysis suggested that the gain in the reading efficiency scores started to stabilize at around Sessions 8 and 9, after which variation in the scores observably diminished (see Figure 1 below). The data set was then divided into two parts—sessions 1-8 vs. sessions 9-16—and the slope coefficient of each part computed. The resulting slope coefficients substantiated the visual inspection; the slope (B = 86.75) of gains in the second half of the training (sessions 9-16) is non-significant at $t = .611$, $p = .564$ as opposed to first half's slope of gains (B = 1170.05), which is significant at $t = 4.091$, $p = .006$.

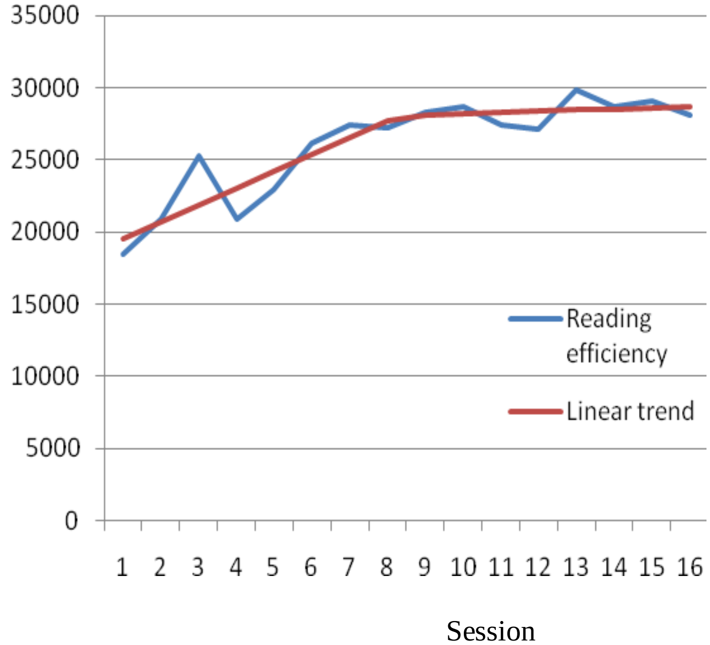


Figure 1. Linear Trend Analysis of reading efficiency scores as a function of sessions

The Linear Trend Analysis results suggest that gains in reading efficiency (i.e., fluency) may be sufficiently achieved from the first eight sessions of training and that additional training sessions may not be necessary for significant gains.

Experimental Study

The study followed a pretest-posttest, control-group, experimental design (Creswell, 2003). See Table 3 below for design notation.

Table 3

The Pretest-Posttest Control-Group Design

Group	Condition
Repeated Readings	R --------------O--------------X--------------O
Wide Reading	R --------------O--------------X--------------O
Vocabulary Study	R --------------O------------------------------O
Comparison Group	N --------------O------------------------------O

Note. R = Random assignment; O = Observation; X = Treatment; N = Non-random assignment

Participants and Setting

Data for the study was collected from struggling readers who had below-college level reading skills and non-struggling readers whose reading achievement was at or above college level. All struggling students were enrolled in a Reading Course at Site 2. Thirteen non-struggling readers provided comparison data for data analysis. Only pretest data from the non-struggling students were used in the analyses conducted. The majority of the non-struggling students were recruited from Site 1 in Spring 2009 (n = 10); the rest of the non-struggling participants (n = 3) were students in the Summer 2009 Reading Course, in which the study was conducted. See table below for study participants' demographic information.

Table 4

Study Participant Characteristics

	Struggling	Non-struggling
Site 1—Spring 2009 (n= 10)		
Age	0	*n* = 10 ; *M* = 22.7; *SD* = 3.97
Male	0	2
Female	0	8
Site 2—Summer 2009 (n=33)		
Age	*n* = 29; *M* = 28.62; *SD* = 8.781	*n* = 3; *M* = 21.00; *SD* = 2.646
Male	7	0
Female	23	3

While for all non-struggling students English was the native language, a variety of languages were spoken by the struggling students: Creole (1), Khmer (1), Gujarati (1), Hindi (1), Telugu (1), Cantonese (1), Portuguese (1), Afrikaans (2), and English (21). None of the students were aware of the hypotheses under investigation. Displayed in the table below is an overview of the study.

Table 5

Overview of Study (Summer 2009)

	May	June	July	August/September
Three sections of Reading Course at Site 2 recruited	X	X		
Pretests administered to 33 students		X		
30 struggling readers received training in RR, WR, or VS for 3 weeks		X	X	
Posttest administered			X	
Data analyzed; results written up			X	X

Note. RR = Repeated Readings; WR = Wide Reading; VS = Vocabulary Study.

Intervention

Three reading classes at Site 2 were recruited with a total enrollment of 33 students. Of the 33 students enrolled in the Reading Course, 30 were identified as struggling readers and three students as non-struggling students based on their performance on the reading comprehension subtest of the *Nelson Denny Reading Test* (Form H; Brown, Fishco, & Hanna, 1993). The struggling readers achieved below the 13th-grade level on ND reading comprehension whereas the non-struggling readers achieved at or above the 13th-grade level. The struggling readers were randomly assigned to one of three conditions (i.e., Repeated Readings, Wide Reading, and Vocabulary Study).

Because the intervention was delivered during participants' class time, a Vocabulary Study condition was designed for the struggling readers (n = 10) assigned to the control condition and for the non-struggling comparison students (n = 3). The Vocabulary Study students individually studied academic words deemed important for college readers to know (Research & Education Association, 2007); they were not exposed to fluency instruction. At each session, a new set of 15 academic words were presented in a binder to the Vocabulary Study group, who (a) perused the list carefully, (b) looked up the words in dictionaries or on the Internet for further study, (c) took a quiz, and (d) filled out a Vocabulary Study Card for words they missed on the quiz. Using the dictionaries provided, the Vocabulary Study students filled out a Vocabulary Card with the spelling, pronunciation and definition of each word they missed on the quiz. They were also required to find an example sentence for the word as well as to write an original sentence of their own. Like the RR and WR groups, the Vocabulary Study students took all pre- and post-test measures as well as the progress monitoring measure of mazes. The Vocabulary Study binders were compiled with (a) a Vocabulary Study sheet, (b) a Vocabulary Quiz (c) the Answer Key, (c) the Progress Graph, (d) Vocabulary Study Cards, and (e) maze tests. At each session a new set of 15 vocabulary words were provided to Vocabulary Study participants.

The study lasted for a maximum of 9 sessions of 30 min in the span of three weeks. The training was incorporated into the regular classroom instruction. During training, arrangements were made for all students to work quietly. While the fluency training students were seated in one section of the classroom to unobtrusively see the large stop-watch screen used for obtaining reading times, the Vocabulary Study students were seated by the computers for further investigation of words on the Internet. All groups completed two weekly maze tests, one prior to Session 4 and the other prior to Session 7. The study began the second week of June and ended the first week of July, 2009, lasting a total of 3 weeks. Posttest measures were administered following the last week of training. The general procedures involved in the intervention program are listed in Table 6 below.

Table 6

Intervention Procedures for Experimental Groups

Condition	Duration	Procedures
Repeated Readings	30 min/session; 3 sessions/week; 3 weeks	Review instructions and progress graph. Read **one passage** on grade level silently. Read the passage **four times** back to back. Plot reading rate for each reading on Progress Chart. Answer 10 questions about the passage. Record comprehension accuracy on Progress Chart. Take a weekly maze test.
Wide Reading	30 min/session; 3 sessions/week; 3 weeks	Review instructions and progress graph. Read **four passages** on grade level silently, **each one time**. Plot reading rate for each passage. Answer three questions (questions 4, 5, 6) about each passage. Record comprehension accuracy on Progress Chart. Take a weekly maze test.
Vocabulary Study	30 min/session; 3 sessions/week; 3 weeks	Study 15 academic words using dictionaries. Take the quiz. Complete a Word Card for unfamiliar words. Take a weekly maze test.

The instructional criteria used with Pilot training students were employed in the intervention study in order to provide instructional challenge and promote students to the next grade-level when they made adequate progress. However, none of the students met the instructional criteria, most likely due to the short duration of the training program (i.e., 9 sessions). It was emphasized to the WR and RR participants that speed and comprehension are the ultimate goals of reading and that they were to answer the comprehension questions as accurately as possible. The Vocabulary Study group was instructed to fill out the Vocabulary Study Cards as completely as possible.

Components of effective fluency intervention.

Several instructional components have been noted for their role in substantial fluency gains by previous research and review studies (Kuhn, Stahl, & Center for the Improvement of Early Reading Achievement, 2000; National Institute of Child Health and Human Development, 2000a; Therrien, 2004b) and thus recommended to be incorporated in fluency instruction. Based on the results of his meta-analysis, Therrien (2004) recommends a combined cue (rate and comprehension) and a fixed number of readings for gains in nontransfer passages. For gains in overall reading rate and comprehension (transfer passages), he recommends using corrective feedback and using a performance criterion (i.e., rereading a text until a preset reading rate is reached). Although the goal of the present study was to test effects on students' higher order comprehension skills (gains in both rate and overall comprehension), a fixed number of readings was implemented for two reasons. First, the comprehension questions accompanying the passages are designed to assess overall comprehension of the passages, with each passage 400 words long; hence enforcing a performance criterion (e.g., reading only the first 100 words) may compromise students' ability to answer all comprehension questions without having read the passage in its entirety. In addition to being cued for comprehension, students were also cued for higher reading rate, a recommendation Therrien makes for non-transfer passages.

Second, the effect sizes (ES) obtained from the meta-analysis suggest that increasing the number of readings may offset the challenges posed by not using a performance criterion. Increasing the number of reading times was accompanied by larger effect sizes in reading rate and comprehension: while a mean fluency ES of .37

was obtained for two readings, an ES of .42 was observed for three readings. The increase was more pronounced for mean comprehension ES from two to three readings: .03 to .49 respectively.

In addition, using charting may augment the ES gains from fixed time readings. Interventions that used charting yielded a mean fluency ES of .57; the mean ES for comprehension was not as large: .11. In this study, participants' reading rate and comprehension scores were plotted on their charts during the training session. The Progress Charts were reviewed by the participants prior to training each session.

Because the training was conducted in this study individually, corrective peer feedback was not provided. Therrien's (2004) review of the literature noted a substantially large effect size of 1.37 for fluency in transfer passages when an adult provided corrective feedback. The analog effect size for the non-transfer passages was not clear-cut; not providing corrective feedback seemed to lead to slightly larger gains. Comprehension was not included in studies assessing the impact of corrective feedback.

An average of 32.5 sessions of Repeated Readings was reported in Therrien's (2004) review. Although nowhere close to this average, particularly in Part 2 with only a maximum of nine sessions, the sessions in this study were longer than the average reported by the National Reading Panel (i.e., 15-30 min) for fluency work embedded in literacy instruction. The longer sessions in this study may have served to compensate for the short duration of the treatment.

Finally, providing challenging reading material was noted as an important component of effective fluency instruction by the authors of the 2000 Center for the Improvement of Early Reading Achievement (CIERA) report (Kuhn et al., 2000). In keeping with the CIERA recommendation, a set of instructional criteria was instituted. Accordingly, students were advanced to harder passages (a) if they read at 400 wpm in a row and achieved 80% comprehension accuracy per session and (b) repeated this achievement over three consecutive sessions.

Materials

Reading materials comprised passages adapted from *Timed Readings*, which is "designed to provide plentiful practice in building reading speed— and comprehension—using graded selections of standard word length" (Spargo & Williston, 1975, p. 7). There are eight levels to the series that cover grades 4-13. In each level, there are 50 passages of a variety of topics of ordinary knowledge to average readers (see Appendix E for a sample passage). Each passage is followed by 10 comprehension questions; five requiring literal comprehension of the passage content and five requiring inferences from the passage. Each passage is 400 words long and has a time limit of a total of five minutes. In its typical application, students record their reading time in minutes and seconds and use a table in the back of the book to convert their reading time to a words-per-minute rate (see Appendix D) and plot the rate on a graph (see Appendix C) for the passage they have read.

Although there is no consensus over the optimal difficulty level for materials to be used in fluency instruction, texts with controlled vocabulary may result in greater gains in reading rate (Hiebert, 2005b) via increased redundancy (Moyer, 1982). The effect of practice is augmented in repeatedly reading a controlled text because it allows little variation in the words and phrases used. At each rereading, words become more familiar and automatic, and this facility leads to greater use of contextual and syntactic cues (Schreiber, 1980).

In Hiebert's Text Elements by Task Model, text difficulty has important consequences on the effects of fluency instruction (Hiebert & Fisher, 2005). Dense texts (i.e., with a higher rate of rare, multisyllabic words mostly found in literary passages) pose a cognitive overload in decoding and retrieving the meaning of rare words. Or, such texts may encourage reliance on context, a strategy that Pikulski and Chard (2005) admonish for taking away students' attention from word processing. By contrast, scaffolded texts are controlled in density with a lower number of rare and long words. In fact, Hiebert and Fisher (2005) identify 73% of the studies reviewed by the report of the National Reading Panel as scaffolded and attribute the moderate effect size (.48) of the report to scaffolded texts used in the studies reviewed by the report. In a more direct investigation of the text effects, a scaffolded-text group of second-graders made greater fluency gains than the literature-based group (Hiebert, 2005b). While the point that Hiebert makes about using scaffolded texts is well taken, scaffolding may take another form in fluency interventions: using an appropriate level of text difficulty might result in larger improvements in students' fluency (Kuhn et al., 2000). In Kuhn and Stahl's review, six out of 11 studies documented gains in treatment groups who read passages at or above their instructional level—texts with no more than 1 in 10 difficult words.

In addition to controlling difficulty for individual texts, difficulty could be controlled for future texts that are in the repertoire of the program. This inter-text manipulation enables transfer of gains to unpracticed passages. An earlier investigation by Rashotte and Torgesen (1985) controlled the words used in each new passage read by learning disabled children in a repeated readings program. Ninety percent of the words were shared across the passages. The results revealed a higher reading rate on new passages after practicing passages with high word overlap than after reading passages with low word overlap.

Because the passages in *Timed Readings* cover scientific topics and social studies, word density was not expected to be high; conversely, consistent with informative prose, words were expected to be repeatedly used to ensure clarity for the reader of what usually is an unfamiliar topic. The caveat suggested by Kuhn and Stahl (2000) was used when the passages were assigned to each student. With students' skills (reading comprehension and decoding) in mind, appropriate passages were assigned from *Timed Readings* at students' grade level. Not only was the use of *Timed Readings* passages expected to expose participants to frequent words, it was also expected to lead to greater gains in reading rate and comprehension since speed and comprehension cues were presented through the use of a reading rate graph and comprehension questions. The following table displays the exposure of study conditions to materials and language input.

Table 7

Exposure to the Materials by Experimental Condition

	Number of Passages Read per Session	Number of Questions Answered per Session	Number of Repeated Readings per session	Number of Passages Plotted per Session	Number of Words Read per Session	Maximum Number of Passages Read	Maximum Number of Words Read
Repeated Readings	1	10	4	1	400	9	3600
Wide Reading	4	12	1	4	1600	36	57600
Vocabulary Study	0	0	0	0	0	0	0

Measurement

Pretest and posttest measures were collected in two different sessions. In one session, paper pencil tests were administered in the following order: the Nelson-Denny Reading Test, maze, MARSI, ART. Computer-based tests of Error Detection, TOWRE, and Reading Span were administered in another session, with their order counterbalanced across participants. The paper and pencil tests were administered in a fixed order due to concerns that statements on the MARSI survey regarding reading behaviors may influence participants' performance on the reading measures if they were to take the reading measures following the MARSI test. Because a similar concern did not exist for the computer-based tests, the order of these measures was counterbalanced at both pretest and posttest. A Latin-square design was used to create possible permutations for the order of computer-based tests. Participants were then randomly assigned to one of three orders derived from the Latin-square design. At posttest, participants were assigned to a different order. Students also completed a demographic survey before taking the computer-based tasks at pretest (see Appendix F for the survey). The demographic survey was used to ascertain basic sample characteristics including age, gender, year in college, ethnicity, parents' educational level, and reading habits.

Reading Course students were tested first either on the paper-n-pencil measures or the computer-based measures by the experimenter and four undergraduate students majoring in Psychology, who served as research assistants. All research assistants satisfactorily completed a training session on all test administration and scoring procedures for each task within the testing battery before taking part in testing and scoring. A tutoring room and participants' classrooms served as the location for testing.

Standardized test protocols were followed during the administration of the *Nelson Denny Reading Te*st and *TOWRE*. Instructions for the Error Detection and Listening Span tasks are provided in the Appendixes G and H, respectively. The *Nelson Denny Reading Test* was used to gather raw, scaled, and grade equivalent data in reading comprehension and vocabulary. The first 60 seconds of the reading comprehension subtest was used to collect the reading rate data of the students while reading silently. The *Test of Word Reading Efficiency* (*TOWRE*) was used to determine participants' isolated word and nonword recognition efficiency. The Reading Span Task was used to measure students' working memory span. Students also completed the Error Detection task, which measured their ability to spontaneously monitor their comprehension of written discourse. All pretest measures were re-administered at posttest except for the reading survey and the ART measure. Table 8 below lists the measures and the points at which they were administered.

Table 8

Measures across Points of Assessment

	pre-test	*training*	*post-test*
ND Reading Test	X		X
TOWRE	X		X
Error Detection	X		X
Reading Span Task	X		X
MARSI	X		X
Maze	X	X	X
ART	X		
Reading survey	X		

Note. ND = Nelson Denny; *TOWRE* = Test of Word Reading Efficiency; MARSI = Metacognitive Awareness of Reading Strategies Inventory; ART = Author Recognition Task.

Error Detection (40 min).

To obtain error detection data, short narrative passages were presented on the computer screen. Materials were adapted from O’Brien and colleagues (Albrecht & O'Brien, 1993; Cook et al., 1998; Hakala & O’Brien, 1995; O'Brien & Albrecht, 1992; O'Brien et al., 1998). Passages began with a two- to three-sentence introduction followed by a section elaborating the characteristics of the main character. The elaborations of the character were either consistent or inconsistent with the target sentence. A filler section followed the elaborations; the filler section ensured that the character elaborations are demoted in passage focus but the story line is maintained. The filler section was either one sentence (local condition) or three sentences long (global condition); thus introducing the distance variable of global or local coherence. It was aimed by including the filler section that characteristics of the story character are purged from the contents of working memory. A target action followed the filler section and described an action that was either consistent or inconsistent with the elaboration of the character. The story concluded with a brief conclusion, which was composed of two to three sentences, followed by one comprehension question. No reference was made to the characteristics of the protagonist in the questions. Passages ranged in length from 20 to 24 lines, with a mean of 22.67 lines. Each line was no longer than 53 characters and ended with a complete word although not necessarily with a complete phrase or sentence.

These manipulations resulted in four different versions of each passage (i.e., global-consistent, global-inconsistent, local-consistent, and local-inconsistent). The different versions were counterbalanced across four material sets. Each set contained 28 passages at pretest and 24 passages at posttest, with 7 passages in each of the four conditions at pretest and 6 passages in each of the four conditions at posttest. Of the 28 passages at pretest, four were randomly removed at posttest due to time constraints. Each subject received a different random order of the passages generated by the DMDX program, which controlled the presentation of the passages and kept a complete record of the latencies and the accuracy of the responses to the comprehension questions about the passages. At pretest, three practice passages were shown on the screen of a laptop computer to familiarize the subjects with the procedures of the Error Detection task. Only one practice passage was used at posttest due both to time constraints and increased familiarity of participants with the task at this point.

A trial consisted of the following events: The experimenter checked the accuracy of the list randomly assigned to the subject. Second, the subject was asked to read along the instructions on the computer screen as the experimenter read them aloud. Following the instructions, the subject was asked to complete the practice trial. Any questions from the subject were answered, and then the subject proceeded to read on the computer screen the passages that made up the experiment.

Subjects were instructed to rest their thumbs on the spacebar. Each trial began with the phrase “Press the spacebar to begin the next passage” presented in the center of the screen. The passages were presented one line at a time, with each key press erasing the current line and presenting the next line. Comprehension time was measured as the time between key presses. At the end of each passage, the word “question” appeared, followed by a comprehension question. To answer the question, subjects pressed either a *yes* or a *no* key. In response to questions answered correctly, the word “correct” appeared as feedback and the word “wrong” appeared for those answered incorrectly.

Since reaction time data do not provide norm-referenced scores, any increase in a treatment group has to be compared to scores from a normally achieving group of readers. Therefore, data from college-level readers were used to provide a source of comparison for the struggling students’ performance in detecting inconsistencies.

ND Reading Test (35 min).

The ND Reading Test was administered to all participants at both pretest and posttest. Alternate forms were used at pretest (Form H) and posttest (Form G). Participants’ gains in comprehension, vocabulary, and silent reading rate (i.e., context-reading fluency) were ascertained from a comparison of the pre-test and post-test scores on the Nelson-Denny measures.

The test was developed by Brown, Fishco, and Hanna (1993) to be used with secondary and post-secondary students. There are two parts in the test: (a) vocabulary and (b) comprehension and reading rate. Students were allotted 15 min to complete the vocabulary part consisting of 80 questions in multiple-choice format. Students are allowed 20 min to complete the comprehension subtest, which consists of seven passages with a total of 38 accompanying comprehension questions. Students read the passages silently and answered literal and inferential multiple-choice questions. The first minute of the comprehension subtest was used to determine the reading rate.

A reliability coefficient of .81 was reported by the developers of the test for the forms G and H; however, no information regarding test-retest reliability of the measure is available.

Reading Span Task (20 min).

This task was administered both at pretest and posttest. The task was first developed by Daneman and Carpenter to measure the working memory capacity of college readers (1980). The measure was designed to be compatible with the Construction and Integration (CI) Model, which assumes that an average reader's working memory span is about three propositions; most readers are not able to accommodate more than three propositions. Therefore, the reading span task provides an index of the number of sentences (propositions) that occupies the working memory.

In the original task, subjects read a series of sentences aloud at their own pace and recalled the last word of each sentence. The task consisted of 60 unrelated sentences that were 13-16 words in length and were presented on index cards. The subjects were required to read each sentence aloud while they were shown the sentences one card at a time. The subjects read all sentences until they saw a blank card, which signaled that they were to recall the last words from each sentence in order. The number of sentences increased from two to six presented in three sets at each difficulty level. The sets presented to the subjects grew increasingly longer until they failed all three sets at a particular level. The level (2-6) at which two of three sets were all correctly recalled was taken as the span of the reader. In the original study, the reading span for 20 participants varied from 2 to 5 with a mean of 3.15 (*SD* = .93).

In the listening span version of the reading span task, subjects had to listen to a sentence and enter a true/false answer. Sentences were presented in five sets each of two, three, four, five, and six sentences. The true-false component was included to ensure that students were processing the sentences and were not concentrating on the last word of the sentences. Subjects listened to the sentence and had 1.5 seconds to enter a true or a false answer before the next sentence was presented. Subjects were stopped if they failed to recall the sentence-final words of all five sets at a particular level. Subjects' span was determined as the level at which they were correct on at least three of five sets of sentences. If they were correct on two out of five sets they were given half point. For example, if a subject recalled three sets at the level of 4 sentences, s/he would be assigned the span of 4. If s/he was correct on only two of the five sets, a span of 3.5 was given.

A meta-analysis by Daneman and Merikle (1996) indicated a weighted reliability estimate of .80 over 473 studies which used the measure and reported reliability scores.

The task therefore is a consistent measure whose application in new settings should not yield deviant scores.

In this study, subjects' storage and processing of verbal stimuli were measured individually at a computer terminal. Subjects read a series of unrelated sentences; the number of sentences in each series increasing from two to five for five sets in each series. Subjects were asked to make validity judgments for each sentence by pressing a *yes* key if the sentence made sense or a *no* key if it did not. When they saw a question mark on the screen, they were to recall the final words from each sentence to a head-mounted microphone.

In this study, total number of words recalled was used to derive a reading span score for each subject as recommended by recent research (Friedman & Miyake, 2005). Friedman and Miyake reports that this method of scoring is more normally distributed, has higher reliability and higher correlations with criterion measures (i.e., reading comprehension) than the traditional span score of the highest set size at which subjects recall correctly sentence-final words.

The Test of Word Reading Efficiency (TOWRE, 5 min).

The Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999) was used to measure participants' context-free word reading fluency and efficiency with decoding. The *TOWRE* was computerized using the *DMDX* software program, which presented the lists of words and nonwords each for the duration of 45 seconds. Participants read aloud the lists as they appeared on the screen to a microphone which recorded their responses until the lists were removed from the screen when the 45 seconds were up. Thus, *TOWRE* was used to measure participants' ability to read words in isolation (i.e., context-free) at both pretest and posttest. Using the *TOWRE* at pretest helped establish a baseline. Pretest results also helped disclose the degree to which the context-free word reading skills determine the level of achievement in reading comprehension and context reading fluency (i.e., the Nelson Denny Silent Reading Rate). Using the *TOWRE* at post-test served to identify treatment effects, if any.

The earliest explanation provided by LaBerge and Samuels (1974) as support for the automaticity theory underscored speed of word recognition, which eventually led to the development of a fluency program (Repeated Readings; Samuels, 1979) to increase word recognition pace. The automaticity theory is supported by research indicating that word recognition is time consuming for poor readers who need more frequent exposures to increase the speed with which they recognize words (Ehri & Wilce, 1983; Reitsma, 1983). According to the automaticity model, a reader cannot execute the processes of word recognition and comprehension simultaneously if her/his lower level skills of word recognition are not automatic. Like the automaticity theory of LaBerge and Samuels, the Verbal Efficiency Theory of Perffetti (1985), postulates that slow reading rate (inefficient word recognition) skills may cause a “bottleneck” which ends up depleting working memory resources of the reader that are necessary to hold large units of text in memory (Shankweiler & Crain, 1986). Therefore, the *TOWRE* was used to ascertain the extent to which poor word reading skills in this study are related to participants' poor comprehension skills, and particularly to their awareness of the state of online comprehension.

TOWRE is a quick measure of two critical word reading skills: to accurately recognize sight words and to quickly decode pronounceable nonwords. There are two subtests to the *TOWRE*: in the Sight Word Efficiency (SWE) subtest real words are presented to subjects who are given 45 seconds to read as many words as they can. In the Phonetic Decoding Efficiency (PDE) subtest, pseudowords are displayed for subjects who are to read aloud as many of pseudowords as possible in 45 seconds. Decoding nonwords accurately and rapidly taps students' ability to use the cipher knowledge (i.e., regular letter-sound rules of English) whereas the sight word test is a measure of students' ability to apply their knowledge of the graphemes for which the regular pronunciation rules are violated (e.g., pint) or more than one pronunciation is possible (/ea/ in beak, steak, area). Both subtests start with monosyllable stimuli which get harder as multisyllable words appear in the rest of the lists. The Sight Word Efficiency subtest starts outs with high frequency words that are assumed to be in the repertoire of most readers' sight vocabulary (e.g., is, up, cat, red, me, he, to, etc.).

There are two alternate forms to each subtest of equal difficulty: Forms A and B. Scores (number of words or nonwords read aloud) can be reported as percentiles, standard scores, and age/grade equivalents with a mean of 100 and a standard deviation of 15. The *TOWRE* manual reports a mean alternate forms reliability coefficient of .90, with the test/retest coefficients ranging from .83 to .96. The Sight Word Efficiency subtest shows high concurrent validity (.80-.94) with other measures of reading ability as is noted in the manual of the test (Torgesen et al., 1999).

For this study, the subtests were computerized when they were administered both at pretest and posttest. The computer software DMDX (Forster & Forster, 2003) developed by Kenneth Forster of the University of Northern Arizona was used to create the interface, which presented the stimuli and recorded students' responses of (non)words read out loud in 45 sec. The number of words correctly read was used to derive the scores from the subtests of sight word and decoding efficiency.

Maze Task (3 min).

The maze test was administered once a week. The maze scores were used to measure gains from fluency instruction at posttest in the number of maze replacements made.

A product of the Classroom-Based Measurement (CBM; Deno, 1985) research tradition, the maze task has gained popularity due to technical and feasibility features that are superior to most other CBM measures, e.g., question answering tests, retell, and the cloze test (Fuchs & Fuchs, 1992). Not only does it have adequate criterion validity (Guthrie, Seifert, Burnham, & Caplan, 1974; Jenkins & Jewell, 1993), it also has sound psychometric qualities which appear to elude other CBM tests such as the retell and cloze methods which are “inadequate as ongoing measures of reading growth” (Fuchs & Fuchs, 1992). Moreover, the maze test’s technical features appear to be similar to those of the mainstream CBM measure of Oral Reading Fluency (ORF), which is widely known to be an accurate measure of growth in reading (Fuchs & Fuchs, 1992). In fact, the maze task has been shown to be correlated with the ORF as strongly as with standardized reading comprehension tests, with coefficients ranging from .80 to .89 for the maze-ORF correlation and from .77 to .86 for the maze-reading comprehension measures (Fuchs & Fuchs, 1992). However, the limited feasibility of the ORF has tainted its popularity and has led to the endorsement of maze by researchers.

Compared to the cloze test and the retell tests, the slope-Standard Error Estimate (SEE) ratio was lower in the maze task, an indication of lower instability in student graphs. Due to low graph instability, it is easier to detect student growth using the maze test, which makes it a technically stronger monitoring test. Further analysis by Fuchs and Fuchs (1992) reveals that the adjusted units of maze slope and SEE are near identical to those of the ORF measure. Thus, the two CBM methods seem to mirror each other in measuring student growth in reading.

The maze task is modified from the cloze test. Except for the first sentence, every seventh word in the passage is deleted and replaced with a 3-word choice. The student fills in the blanks with the correct alternatives in three minutes. In creating distracter word choices for the maze test, care is exercised to select words that (a) are of similar length to the correct replacement, (b) do not fit semantically with passage context and (c) do not share phonological and orthographic likeness. In addition, distracters should be familiar enough that they are not taken as nonsense words. According to Fuchs and Fuchs (1992), correct replacements should be able to be made within one or two sentences; distracters that require students to read further ahead in the text should not be used.

The maze task has been described as a global measure of reading, requiring decoding, fluency, and comprehension (see Fuchs, Fuchs, Hamlett, & Ferguson, 1992; Williams et al., in print). In this respect, like ORF, the maze task represents not only word level processes in reading but also "processing meaningful connections within and between sentences, relating text meaning to prior information, and making inferences to supply missing information" (Fuchs, Fuchs, Hosp, & Jenkins, 2001, p. 240). As a reflection of the complex orchestration of the cognitive processes that take place during reading, the maze task therefore is assumed to incorporate both word level recoding and text level comprehension skills all at the same time. The sensitivity of the maze procedure to measuring global reading skills was demonstrated in a study of infusing ongoing reading instruction with specific versus generic oral reading fluency instruction in middle school students with learning disabilities or reading difficulties (Allinder et al., 2001). Students in the specific strategy group either self-selected the strategies or were assigned a strategy by their teacher. Specific oral reading strategies included reading with inflection; not adding words, pausing at periods and commas; self-monitoring for accuracy; reading at an appropriate pace; watching for word endings; and tracking with finger. Students in the generic strategy group were told to "do their best." The results revealed that while the two groups did not differ on the posttest measure of a standardized reading comprehension test, the specific-strategies group outperformed the generic group in maze slopes.

Because the maze test is a global measure of reading, the maze test circumvents the problem that troubled earlier fluency measures. The fluency components of DIBELS were found to "mispredict reading performance on other assessments of [reading ability] much of the time, and at best is a measure of who reads quickly without regard to whether the reader comprehends what is read" (Pressley, Hilden, & Shankland, 2006, p. 2), because students in DIBELS are cued solely for speed, not for comprehension. Samuels cautions therefore that "what we need…are tests that mimic fluent reading, that demand simultaneous decoding and comprehension. In order to do that, the researcher must inform students that as soon as the oral reading is done, the student will be asked comprehension questions" (Samuels, 2007, p. 565).

Maze tests used in this study were selected from a collection of eighth-grade maze tests downloaded from the AimsWeb website (www.aimsweb.com). Keeping the maze passages constant helped keep difficulty level constant throughout the study, thus enabling measurement of growth (Deno & Marston, 2006).

Author Recognition Task (ART).

Developed first by Stanovich and West (Stanovich & West, 1989) with adult readers, the ART has been a consistent and robust measure of reading exposure in different populations (Cunningham & Stanovich, 1991, 1997, 1998; Stanovich & Cunningham, 1993). The test is comprised of a number of real author names and a number of foils. Students are asked to check the names that they are sure are real authors. They are warned against guessing. A score is derived by subtracting wrong answers from the correct answers. Recently, Acheson, Wells, and MacDonald (2008) have updated the ART, which is attached in Appendix I.

In this study, the ART was administered at pretest to yield an estimate of subjects' print exposure.

Metacognitive Awareness of Reading Strategies Inventory (MARSI, 3 min).

The MARSI survey was developed by Mokhtari & Reichard (2002) to measure awareness and use of reading strategies in 6th- through 12th-grade students reading academic texts. Different from existing measures of metacognitive awareness, the MARSI survey has (a) a larger number of items per scale, (b) sound psychometric properties and (c) strong construct validity. The measure is comprised of three scales: (a) Global Reading Strategies; (b) Problem-Solving Strategies; and (c) Support Reading Strategies. The 13 Global Reading Strategies items are statements of strategies that describe a "global analysis of text" (Mokhtari & Reichard, 2002, p. 252) such as "I decide what to read closely and what to ignore;" "I think about what I know to help me understand what I read;" and "I have a purpose in mind when I read." There are eight items in the second scale, Problem-Solving Strategies, that describe reading strategies for resolving reading difficulties while reading text such as "When the text becomes difficult, I reread to increase my understanding;" and "I adjust my reading speed according to what I read." Nine items in the third scale, Support Reading Strategies, describe using strategies that support reading comprehension through the use of various materials or aids. Items in this scale include "I take notes while reading" and "I underline or circle information in the text."

The MARSI survey was administered both at pretest and posttest (The MARSI survey is attached in Appendix J).

Threats to Internal Validity

To the extent that the "changes observed in the dependent variable are due to the effect of the independent variable, not to some other unintended variables" (Mertens, 1998, p. 64) internal validity is established in an experimental research study. Threats to the internal validity of the study were minimized by recruiting a control group who experienced the same study conditions throughout the study.

Extraneous variables, which are best controlled for in laboratory settings, are a major concern for research conducted in educational settings, for educational settings are fraught with threats to internal validity of the research, rendering the direct causal attributions to the independent variable confounded. The strict implementation of the intervention instructions helped ensure that a functional relation is established between the dependent and independent variables and that this relation is not due to extraneous variables emanating from implementation variability of the training procedures (Kennedy, 2005). Step-by-step instructions were written for each condition and included in the participants' folders (see Appendix K). Every effort was made for the intervention procedures to be followed strictly by the participants, who reviewed the instructions to their respective training conditions each session prior to training.

History effects arise when there is strong reason to believe that experimental effects may be attributed to events that occur outside the control of the researcher. Extraneous events that may lead to experimental change include health issues, lack of sleep, supplemental tutoring sessions, and the like (Kennedy, 2005). As regards the current study, timing of midterms and final exams may have threatened the attributability of the experimental effects to the fluency training. Students may have been compelled to do more reading than they usually do when studying for midterm and final exams; doing so may have given them extra practice and additional exposure to reading materials. However, these effects were balanced out by having a control group, who were exposed to the same academic requirements during the study as the Wide Reading and Repeated Readings students but not the fluency treatment (Mertens, 1998).

Internal validity could also be threatened by testing effects. Testing effects are particularly adverse in studies using both pre- and posttests. Similarities between the pre- and post-test measures may have led to test-wiseness among the study participants, who may have become sensitized to test procedures. In the current study alternate forms were used for all assessment measures except for the ERDE task. Because an alternate set of passages did not exist for the posttest measurement of ERDE, the pretest passages were reused. Testing effects were sought be mitigated in two specific ways. First, the participants were assigned to a different list of the ERDE task at posttest than the pretest. It was expected that assigned to a different list of ERDE passages, participants would be exposed to a different version of the passages than the pretest. For example, if a given participant was assigned to List 1 at pretest, s/he was assigned to List 4 at posttest and was presented with the global-inconsistent version of the Bill passage, which was presented as a -local-consistent passage at pretest. Second, the comprehension questions participants answered about ERDE passages were rewritten. Because the participants were instructed to “read the passages at their own pace and answer the comprehension questions as accurately as possible,” the comprehension question was thought to be the most salient part of the ERDE task for a participant. It was expected that using a different comprehension question would greatly reduce recall of passage details. Therefore, to minimize memory effects, a new comprehension question was written for each ERDE passage.

In addition, there were several occasions on which the portable laptop computers used for data collection malfunctioned during the administration of the pre-test and post-test assessments. Software glitches required re-administration of the computer-based tasks for the affected participants. One student’s pretest responses to the TOWRE were not recorded by the computer due possibly to another software program occupying the audio system of the computer. Three students were re-administered the ERDE and RSPAN tasks. A memory advantage may have accrued for participants who were re-administered the computer-based tasks, possibly improving their performance.

As deleterious to internal validity are maturation effects. The effects of an intervention program may be confounded by biological, cognitive, and emotional changes that may occur during the course of a study in participants. Maturational changes may include “becoming stronger, more coordinated, or tired as the study progresses” (Mertens, 1998, p. 65). However, having a control group balances out the effects of maturational change because the control group students also experience similar change, but not the treatment.

Procedural Integrity

A detailed, step-by-step procedural integrity check-list was written for the experimenter to follow during the implementation of the intervention procedures. A set of instructions were written for each of the Repeated Readings, Wide Reading, and Vocabulary Study Group conditions, which are attached in the Appendix L. The experimenter made every effort to strictly follow these directions. The instructions for the experimenter ensured that (a) materials were present during treatment, (b) intervention instructions were reviewed by the participants prior to training each session, (c) the participants followed the proper sequence of their respective intervention procedures, and (d) the treatment conditions were implemented correctly. The instructions for the experimenter also specified the unique features of the fluency training conditions and whether or not they were implemented correctly.

Summary

A fluency intervention program was implemented in Summer 2009 to remediate fluency deficits in a group of struggling undergraduate students. With their reading ability level determined, subjects reading below-college level were randomly placed in one of three training conditions: Repeated Readings, Wide Reading, and Vocabulary

Study Group. Because the study was conducted with college readers in three sections of a Reading Course offered at a community college, a Vocabulary Study program was designed for students who served as the control group in this study. The Repeated Readings group read a grade-level passage four times back to back; the Wide Reading group read four different passages per session each once whereas students in the Vocabulary Study condition studied academic vocabulary words without engaging in any connected text reading.

The fluency intervention program was piloted in Spring 2009 with struggling college readers prior to its implementation in Summer 2009. The Pilot data from six students who completed at least 16 sessions were analyzed to estimate the dosage of fluency intervention to implement with Summer 2009 students. A Linear Trend Analysis indicated that gains in reading efficiency of pilot students start a stable trend at around Sessions 8 and 9. In light of this result, the Summer 2009 intervention was capped at 9 sessions. Treatment effects were investigated on reading comprehension, vocabulary knowledge, reading rate, verbal working memory efficiency, and ability to construct coherent mental models in students who participated in the training in Summer 2009.

CHAPTER FIVE

RESULTS

This section reports on the results of the statistical analyses conducted on data from pretest and posttest measures. The analyses were performed to answer the following research questions:

(a) How do the study groups compare at pretest on the reading ability measures?

(b) Does fluency training result in significant gains on measures of reading comprehension, vocabulary, silent reading rate, maze, RSPAN, TOWRE SWE and TOWRE PDE? Which fluency program (RR vs. WR) leads to greater gains?

(c) How do the study groups compare at posttest on the reading ability measures?

(d) Does fluency training raise subjects' sensitivity to textual inconsistencies? Which fluency program (RR vs. WR) leads to greater gains?

Screening Data

Prior to analyses, data sets were screened for missing values and non-normality. A discussion of the methods undertaken to deal with missing values and non-normal distributions is presented below before the results of appropriate statistics conducted on the DVs are reported.

Missing Data

Due to computer malfunctions, several students' responses were not recorded on the computer-based tests of TOWRE and RSPAN. Data from these students were treated as missing data, which were replaced by the group mean. Mean substitution is described by Tabachnick and Fidell as a "popular way to estimate missing value" (Tabachnick & Fidell, 2005, p. 67). An advantage to using the mean for missing values is the conservativeness of the procedure: "the mean for the distribution as a whole does not change and the researcher is not required to guess at missing values" (p. 67). Although this procedure may lower the variance of a variable, Tabachnick and Fidell suggest using the group mean as a compromise instead of the overall mean. This suggestion was taken in imputing values for missing data. Missing values for cases were replaced by the group's mean score on that variable. For example, a Repeated Readings student's missing values on the Reading Span task at pretest was replaced by the mean Reading Span score of the Repeated Readings group at pretest.

A total of seven missing values were substituted by group mean values at pretest, affecting 2 Repeated Readings and 2 Vocabulary Study subjects. Mean group scores were imputed for 3 missing values on Reading Span, for 2 missing values on TOWRE PDE, and for 2 missing values on the TOWRE SWE tests. Eight values were missing at posttest involving 2 Repeated Readings, 1 Wide Reading, and 1 Vocabulary Study students. Mean substitution resulted in replacing the group mean for 4 Reading Span, 2 TOWRE SWE, and 2 TOWRE PDE missing values.

Distributional Normality of Dependent Variables

Graphical and numerical methods were conducted to test normality of the distributions of scores on the dependent variables. Summary statistics such as skewness and kurtosis were obtained from numerical methods, and statistical theory-driven tests of normality were conducted. Skewness is a measure of dispersion in the distribution. It measures the degree to which data values deviate from the mean to either the left tail of the distribution (positive skew) or the right tail (negative distribution). A non-zero skew score is also an indication of the direction of the asymmetry; a positive skew score means the data is positively skewed while a negative score indicates the data are piled towards the right end of the distribution away from the mean. A zero score indicates no skew in the data set. Kurtosis, another dispersion measure, is a measure of the “peakedness” or flatness in the data relative to a normal distribution. Highly kurtotic data sets are characterized by a swarm of data peaked around the mean with short tails. On the other hand, a flat top and long tails characterize a data set with low kurtosis.

In addition to numerical (i.e., skewness, kurtosis) and graphical (e.g., box plots, histograms) methods, which provided objective and intuitive ways of examining normality in the data respectively, the Shapiro-Wilk (*W*) statistic was used for testing normality. The *W* statistic is recommended (Park, 2008) for samples sizes greater than or equal to 7 and less than or equal to 2,000. The *W* is reported as a positive number, less than or equal to one. A *W* score close to one indicates a normal distribution of data. For example, according to the Shapiro-Wilk test, the pretest scores on the Nelson Denny reading comprehension test are normally distributed, W=.960, df= 43, p<.140 (see Figure 2 below). In other words the null hypothesis of normality is not rejected at the .05 level.

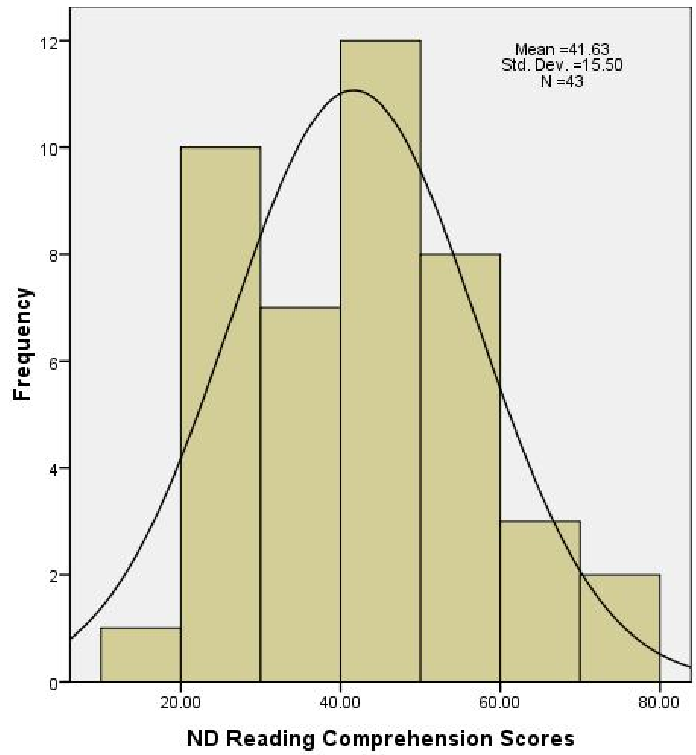


Figure 2. Distribution of pretest ND Reading Comprehension scores.

The W was computed for all dependent variables in the study and listed in Appendix M. Data sets with significant ($p < .05$) W values are marked with an asterisk indicating that the group is not normally distributed.

Data Transformations

Non-normal distributions were detected in the data set: (a) Repeated Readings Pretest ND Silent Reading Rate, $W = .601$, $df = 11$, $p < .001$, (b) Repeated Readings

Author Recognition Test (ART), $W = .855$, $df = 11$, $p = .049$, (c) Wide Reading Posttest ND Silent Reading Rate, $W = .834$, $df = 9$, $p = .049$, (d) Vocabulary Study Group Posttest Reading Span, $W = .827$, $df = 10$, $p = .031$. This was not an unexpected outcome given the nature of the data. A cutoff score of 13th-grade level was utilized to define groups as skilled and less skilled readers. Scores from a group of subjects who fall into the range of a set of cutoff scores may not readily follow a normal distribution. However, attempts were made to normalize distributions that seem to be affected by the presence of outliers. Tabachnick and Fidell (2005) suggest two methods to reduce the impact of outliers: (a) transform variables or (b) change scores. First, data transformation was attempted to deal with non-normal distributions. However, because transformation methods employed either did not improve normality of distributions or affected all groups on the offending variable and thus altered previously normal distributions to non-normal, the second recommendation by Tabachnick and Fidell was taken to normalize the data sets with large dispersion.

For example, Vocabulary Study Posttest Reading Span scores were found to be non-normally distributed, $W = .827$, $df = 10$, $p = .031$, with moderate skew (-1.548) and kurtosis (2.292) as shown in Figure 3 below.

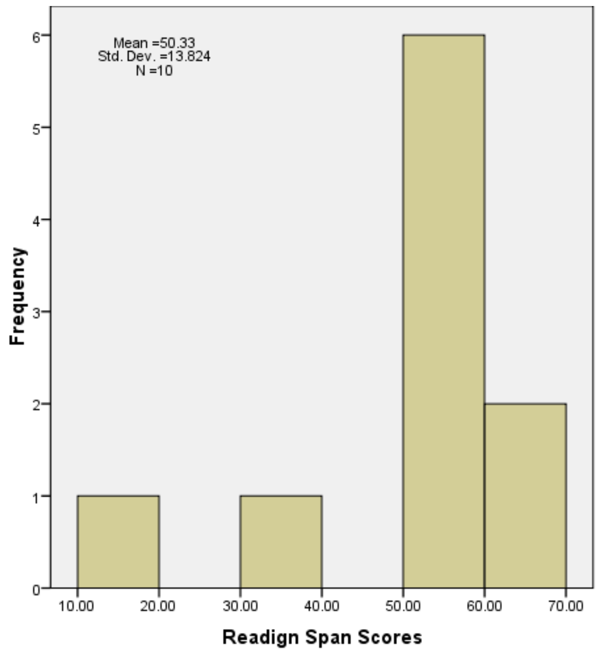


Figure 3. Distribution of Vocabulary Study Group's posttest Reading Span scores.

To improve normality in this group's data, the square root method was employed, followed by the log and the inverse methods. The first two methods are recommended for data that exhibit moderate right skewness. The inverse method is useful for removing severe positive skewness (Tabachnick & Fidell, 2005).

Taking the square root and the log of the data set did not remove the skew in the data. The skewness was increased to -1.840 and the kurtosis to 3.361 following the square root method. The log method resulted in larger skewness at -2.117 and further increased the kurtosis to 6.909. Neither method transformed the data to (near) normal statistically (*Wsquare root*= .777, *df*= 10, *p*= .008; *Wlog*= .723, *df*= 10, *p*= .004). To improve the normalization of dispersion in the data, the inverse method was used next, which resulted in relatively more skew. With the inverse method, the skew in the data was increased to 2.582. The *W* was significant at .612 (*p*<.001) following the inverse method, indicating that the comparison group's data set was still non-normally distributed (see Figure 4 below).

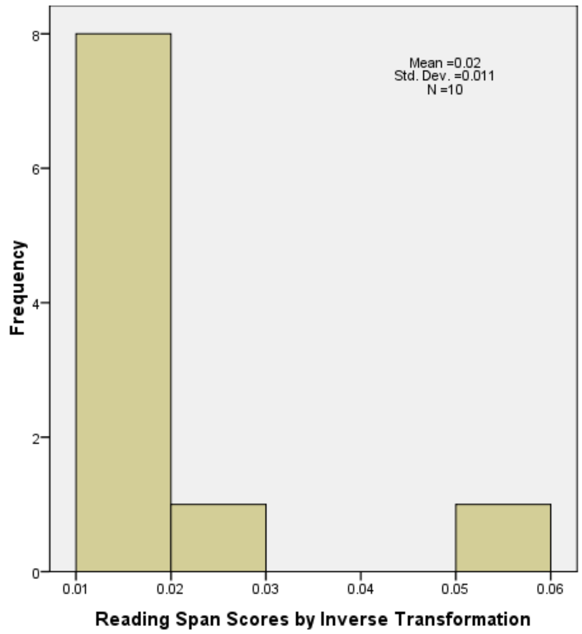


Figure 4. Distribution of Vocabulary Study Group's posttest Reading Span scores following transformation.

Following the second suggestion by Tabachnick and Fidell (2005), outliers were sought within each group whose data set was found to be non-normally distributed. Tabachnick and Fidell define outliers as "cases with very large standardized scores, *z* scores, on one or more variables, that are disconnected from the other *z* scores" (p. 73). Potential outliers are assumed to have "a *z* value of 3.29 or larger" (Tabachnick & Fidell, 2005, p. 73), which may by chance appear in data sets with a large sample size (e.g., *n* >100). For data with a smaller sample size, more conservative *z* values are used to identify the outliers; Stevens (1999) suggests using a *z* score of 2.5 to consider a data point as an outlier. According to Shiffler (1988 cited in Stevens, 1999, p. 14) for a data set with ten subjects, any data point with a *z* value larger than 2.846 is an outlier, and for a sample size of 11 any *z* value in excess of 3.015 is an outlier.

Lack of normality was detected by the *W* test in the pretest ND Silent Reading Rate scores of the Repeated Readings group, $W = .601$, $df = 11$, $p < .001$. Inspection of the group box plot and the *z* values revealed an outlying score for subject *pr_gpc_38*. The score for this subject (i.e., 566 wpm) fell within 4.208 standard deviations of the mean. A reading rate score of 566 wpm, which is beyond the typical reading rate (i.e., 330 wmp) for college readers (Carver, 1990), appeared abnormal for a less skilled reader. Therefore, the outlying score was replaced with a more accurate estimate of reading rate (i.e., 244.89 wpm) for this subject. This score was obtained from the subject's first day of Repeated Readings training. Following the modification of the outlying score, the pretest data distribution for the Repeated Readings group became more near normal, $W = .937$, $df = 11$, $p = .489$.

Because similar ways of estimating a more accurate score for outliers on other non-normally distributed group data sets (i.e., *b*, *c*, & *d*) were not available, a suggestion by Tabachnick and Fidell (2005) was utilized to modify the outlying data values in these offended data distributions. Their suggestion involves "assigning the outlying case(s) a raw score on the offending variable that is one unit larger (or smaller) than the next most extreme score in the distribution" (p. 77). Wide Reading posttest ND Silent Reading Rate data set was subjected to this modification by changing the outlying subject's (i.e., *pr_gpc_40*) posttest ND Silent Reading Rate score. This subject's original score 398 wpm lies 2.642 standard deviations from the mean of the group. The score was changed to 356 wpm, which is one unit larger than the next most extreme score of 355 wpm. The normality of the group's data distribution, however, was not affected by this modification, $W = .829$, $df = 9$, $p = .044$. As a result, no modifications were made on non-normal distributions of *b*, *c* and *d*. Descriptive and normality statistics of all dependent variables (taken at pretest and posttest) are provided in Appendix M; listed in Appendix N are box plot distributions of DVs.

Data Analysis

After screening, the data were analyzed using various statistical tests. The results of these tests and the discussion of the findings will be related to the research questions posed.

Research Question A: How Do the Study Groups Compare at Pretest on Reading Measures?

A one-way Analysis of Variance (ANOVA) was performed to test mean differences among all four groups on each pretest measure of Reading Comprehension, Vocabulary, Silent Reading Rate, MARSI, maze, RSPAN, TOWRE SWE, and TOWRE PDE. ANOVA is appropriate for situations in which more than two groups are simultaneously compared on a dependent variable (Stevens, 1999). The statistic was used to test the null hypothesis that the population means of Repeated Readings, Wide Reading, Vocabulary Study and the Comparison Group are equal on pretest measures.

ANOVA is a robust statistical method against violations of normality and unequal variances with equal or quasi-equal group sizes (Stevens, 1999, p. 76). Assumptions of ANOVA were satisfied for all dependent variables for comparing groups at pretest. First, all pretest data sets were found to be normally distributed (see Appendix M) except for two, which were characterized by slight to moderate departure from normality: Repeated Readings Group ART data set (skew= 1.283; kurtosis= 3.877) and Comparison Group pretest maze data set (skew= -.641; kurtosis= -1.346). Second, except for the ART measure, all dependent variables were found to have equal variance, as determined by the Levene's test of Homogeneity of Variance. The Levene test was not significant (at $p < .05$) for all dependent variables except for ART, $p = .001$. A significant Levene's Test statistic indicates that the *k* variances are significantly different. A non-significant Levene's Test, on the other hand, means that the variances are homogeneous.

The final assumption that all observations be taken independently of each other was considered tenable since the assessment measures were individually administered. Because no interaction was involved among the study participants, the observations were not thought to influence each other (Glass & Hopkins, 1984 cited in Stevens, 1999, p. 78).

The one-way ANOVA resulted in significant overall differences on all pretest measures except for MARSI ($F[3,39] = 1.036$, $p = .387$) and TOWRE PDE ($F[3,39] = 2.474$, $p = .076$). In other words, there were at least two groups with significantly different means on Reading Comprehension (F[3,39] = 22.623, p < .001), Vocabulary (F[3,39] = 13.548, p < .001), Silent Reading Rate (F[3,39] = 9.464, p < .001), ART (F[3,39] = 12.506, p < .001), maze (F[3,39] = 9.257, p < .001), RSPAN (F[3,39] = 3.156, p < .05), and TOWRE SWE (F[3,39] = 4.150, p < .05), but not on MARSI and TOWRE PDE.

To determine where the significant differences lie, the Tukey post-hoc procedure was conducted on each pretest measure with overall significant differences. The Tukey "provides a nice balance in terms of controlling on both Type I and Type II errors, while focusing on meaningful, easily interpreted comparisons" (Stevens, 1999, p. 86). The Tukey procedure revealed that the overall differences were due to the Comparison Group, who outperformed the RR, WR, and VS groups, all at $p<.05$, on Reading Comprehension, Vocabulary, Silent Reading Rate and ART. The three poor reader groups were statistically comparable on these measures. On maze, the Comparison Group was statistically comparable to the WR group but did better than the RR and VS groups. On RSPAN, the Comparison and RR groups differed significantly in the mean number of words recalled ($p < .05$). On TOWRE SWE only the Comparison and VS groups differed in the number of sight words read aloud correctly within 45 seconds, $p = .006$. Other comparisons were not statistically significant at $p < .05$. The Comparison Group's advantage on the reading measures over the poor reader groups is evidence that using a criterion of reading at college level sufficiently demarcated the groups in terms of their reading ability in this study.

The lack of clear-cut group differences on TOWRE SWE, RSPAN, and the findings of non-significant differences on TOWRE PDE and MARSI are surprising. With significantly higher reading comprehension ability, vocabulary knowledge and silent reading rate, the Comparison Group of skilled readers were also expected to read significantly more nonwords within 45 seconds than all groups of poor readers (i.e., RR, WR, VS). Moreover, their advantage on TOWRE SWE was not expected to be limited to the VS group only; they were expected to perform significantly better than the poor readers in the WR and RR groups as well. By the same token, they were expected to recall a significantly greater number of words than all groups of unskilled readers on the RSPAN task, not only the RR group.

These findings appear contradictory to findings from various studies and models that assume a strong link between decoding and comprehension (Bell & Perfetti, 1994; Bruck, 1988, 1990; Cunningham, Stanovich, & Wilson, 1990; Stanovich, 1981). Children's and adults' ability to read words and nonwords with speed is indicative of the quality of their lexical processes: accessing phonological, semantic, and syntactic representations of words in the lexicon. The processes involved in single word/nonword reading also include sublexical processes and sound-to-symbol mappings as well as processes of accessing representations of low frequency and irregular/regular words. For example, the lexical errors committed by a college student reading at the 19th percentile on the *Nelson Denny Test* were limited to low-frequency (hard) words that occurred in a challenging passage (the Holocaust) with vocabulary "a step above the mundane" (Perfetti, Marron, & Foltz, 1996, p. 151). The subject's lexical errors accompanied an observed difficulty to derive a correct understanding of the passage, as was evidenced in her verbal protocol. On a simpler text that followed this passage, the same subject made fewer word reading errors but showed similar difficulty constructing a coherent mental model.

Furthermore, these findings suggest a dissociation of reading skills at higher grade levels where background knowledge and language comprehension play a larger role in determining reading ability (Gough et al., 1996). Students' experiences with print and exposure to print materials may serve to promote their reading ability (e.g., reading comprehension, vocabulary knowledge, silent reading rate) even with relatively low word recognition skills. For example, the Comparison Group of skilled readers, who recognized significantly more popular authors on the ART survey (M = 12.23) than the poor readers in the training groups (M = 3.73), also achieved an average of 14.8th-grade level on the ND reading comprehension subtest; an average of 14th-grade level on vocabulary; and an average of 275.62 words per minute on the ND Silent Reading Rate subtest. By contrast, their overall achievement on TOWRE SWE was at the 9.8th-grade level.

Poor readers, on the other hand, show a tendency of relying on their knowledge of sight words. The grade level achievement of the training groups (RR, WR, VS) in reading comprehension and vocabulary is relatively matched to their achievement on TOWRE SWE. The RR group, who achieved 8.1th-grade level reading comprehension and 9.5th-grade level vocabulary, scored at the 8th-grade level on sight word efficiency. The TOWRE SWE achievement of the WR group, who achieved at 8.7th-grade level on reading comprehension and at the 9.5th-grade level on vocabulary, is at the 9.8-th grade level. The VS group's reading comprehension (9.2th) and vocabulary (9.3th) grade-level achievement is about 3 grade levels greater than that of their sight word efficiency (6th).

The Comparison Group students' greater achievement in higher level reading skills appears to be due to their advantage of 2 grade levels in decoding efficiency over the training groups of poor readers. The students in the Comparison Group achieved an average of 5.6th-grade level on the TOWRE PDE subtest while the RR and WR groups each achieved at the 3.6-th grade level and the VS group at the 3.4-th grade level (see Table 9 below).

Table 9

Groups' Grade Levels and Means on Reading Comprehension, Vocabulary, TOWRE SWE and TOWRE PDE

	NDRC		NDVOC		TOWRE SWE		TOWRE PDE	
	Mean	GL	Mean	GL	Mean	GL	Mean	GL
RR	31.27	8.1	38.55	9.5	81.4	8	27.4	3.6
WR	34	8.7	37.67	9.5	88.46	9.8	27.89	3.6
VS	36	9.2	33.4	9.3	74.67	6	25.78	3.4
CG	60	14.8	55.85	14	88.27	9.8	38.31	5.6

Note. NDRC = Nelson Deny Reading Comprehension; NDVOC = Nelson Denny Vocabulary; TOWRE = Test of Word Reading Efficiency Sight Word Efficiency (SWE); Phonemic Decoding Efficiency (PDE); GL = Grade Level; RR = Repeated Readings; WR = Wide Reading; VS = Vocabulary Study; and CG = Comparison Group.

Poor readers are known to be deficient in applying the English Grapheme-Phoneme Correspondence (GPC) rules to recode non-words, such as those found on the TOWRE PDE subtest (National Institute of Child Health and Human Development, 2000a; Shankweiler et al., 1999; Share, 1995). Nevertheless, a 5.6th-grade level mastery of the English GPC rules, which was demonstrated by the skilled readers in this study, may be sufficient for adult readers to learn vocabulary and comprehend texts at college level. As was suggested by Jenkins, Fuchs, van den Broek, Espin, and Deno, "further improvement in [word recognition] may have less effect on context fluency, and comprehension skills become a stronger determinant" after word reading efficiency reaches a certain level (2003, p. 726). It may be that the 5.6th-grade level decoding skill is the lower limit for achieving college level reading comprehension with an adequate store of sight words and background knowledge.

RSPAN results may make more sense when the results from the TOWRE SWE and TOWRE PDE are taken into account. Only the Comparison and RR groups differed significantly in the number of words recalled; other pairwise comparisons were not significant. Similarly, the groups exhibited commensurate skills in reading words and nonwords; only one pairwise comparison was significant on the TOWRE SWE subtest (CG vs. VS) and no significant group differences were observed on TOWRE PDE. Efficiency in reading words and nonwords may underlie one's efficiency on the RSPAN task, in which subjects are required to read sentences, make a true/untrue judgment, and recall the final words from each of the sentences in the trial. Since the Daneman and Carpenter (1980) landmark article, which pitted processing efficiency against short-term memory, research on the issue has favored efficient processing in the domain of language as it relates to reading (Daneman & Tardif, 1987). In other words, the (non-)word reading skill of the groups may constrain their performance on the working memory measure of RSPAN, which requires processing and storage of verbal input.

Finally, the MARSI findings indicate that awareness and use of reading behaviors and habits may not be a discriminating factor between skilled and less skilled readers. Poor readers may self-report observing certain behaviors during reading that they may not naturally engage in. Despite their relatively low reading achievement, the unskilled readers in this study reported engaging in cognitive behaviors and using metacognitive reading strategies as often as their skilled peers. These findings corroborate previous reports of social desirability effect, a tendency to overreport, for less skilled readers (Stanovich & West, 1989).

In summary, pretest data analyses revealed that (a) the Comparison Group achieved significantly higher than the RR, WR, and VS groups on Reading Comprehension, Vocabulary, Silent Reading Rate and ART while the RR, WR and VS groups did not differ from one another; (b) only the CG and the RR groups were significantly different on RSPAN, (c) only the CG and VS groups differed significantly on TOWRE SWE, (d) the CG group outperformed the RR and VS group on the pretest maze test while achieving commensurately with the WR group, and (e) no reliable differences were observed on TOWRE PDE and MARSI. While the significant group differences on reading comprehension, vocabulary, ART, and silent reading rate verify the method of defining students as achieving at or below college level, lack of clear cut differences on the processing measures adduce further evidence to the role of language comprehension in determining reading ability at higher grade levels. The fact that unskilled readers reported being aware and using reading strategies as often as the skilled readers reiterate the social desirability tendency widely observed in unskilled readers.

Research Question B: Does Fluency Training Result in Significant Gains on the Reading Measures? Which Fluency Program (RR vs. WR) Leads to Greater Gains?

To identify differential gains, if any, by groups from pretest to posttest, significant interaction effects between *time* and *group* variables were sought by computing separate Repeated Measures ANOVAs on each of the dependent variables, i.e., reading comprehension, vocabulary, silent reading rate, MARSI, maze, RSPAN, *TOWRE SWE*, and *TOWRE PDE*. To test group-specific training gains, separate paired-samples *t*-tests were conducted on the pretest and posttest data of the training groups (i.e., RR, WR, VS) on measures, for which the Repeated Measures ANOVA resulted in a significant *time* main effect. The paired-samples *t*-test statistic was utilized to test whether the *time 1-time 2* difference score is “greater than expected by chance alone” (Stockburger, 2001). The analysis tested the null hypothesis that the average difference between *time 1* and *time 2* scores would be zero ($\mu_D = 0$) if an infinite number of subjects participated. What ensue are the results of the Repeated Measures ANOVAs and paired-samples *t*-tests conducted.

Reading comprehension.

No significant effects were observed on the ND Reading Comprehension scores for *time* $F(1,27) = 2.198$, $p = .150$; for *group* $F(2,27) = .577$, $p = .568$; and for *time by group* $F(2,27) = 1.076$, $p = .355$.

Vocabulary.

Only a significant *time* main effect was observed, $F(1,27)= 16.145$, $p < .001$, $\eta^2 = .374$. Overall time 2 vocabulary performance ($M = 41.832$; $SE = 1.653$) was significantly greater than time 1 performance ($M = 36.537$; $SE = 1.375$) across all groups. Neither the *group* main effect nor the *group by time* interaction effect was significant, $F(2,27) = .826$, $p = .448$; $F(2,27) = .431$, $p = .654$ respectively.

The RR students answered 3.73 more vocabulary items correct on the *ND* Vocabulary subtest at posttest ($M = 42.27$) than at pretest ($M = 38.55$). The difference in the means was not significant, $t(10) = -1.818$, $p = .099$. Unlike the Repeated Readings students, students in the WR group achieved a statistically significantly, $t(8) = -2.399$, $p = .043$, $dz = 0.79$, higher vocabulary score at posttest ($M = 43.22$) than at pretest ($M = 37.67$). The largest gain in vocabulary knowledge was achieved by the Vocabulary Study group, who answered 6.6 more vocabulary items correct at posttest ($M = 40$) than at pretest ($M = 33.4$). This gain was statistically significant at $t(9) = -2.674$, $p < .05$, $dz = 0.85$. Mean gain scores by groups in vocabulary are displayed by the following figure.

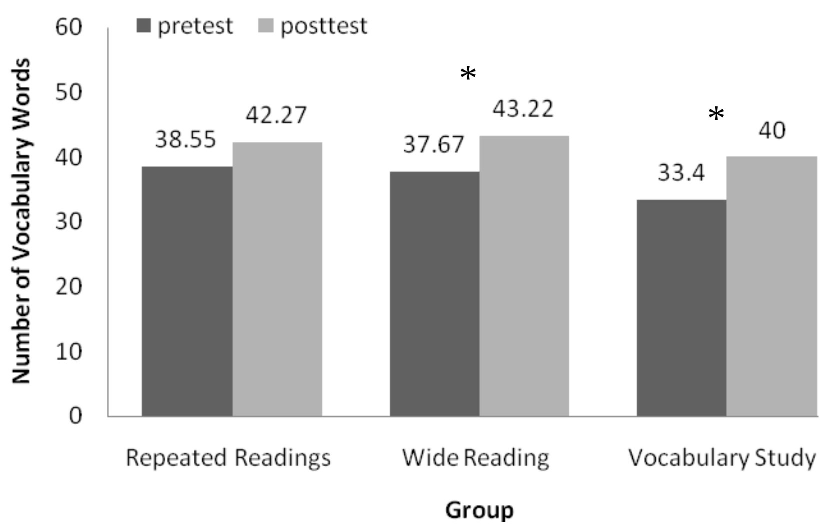


Figure 5. Mean vocabulary gains by training groups.

* $p < .05$.

Silent reading rate.

A similar pattern of results was observed in groups' reading rate performance from time 1 to time 2. Overall silent reading rate at time 2 ($M = 226.27$; $SE = 11.99$) was significantly greater than silent reading rate achieved by all groups at time 1 ($M = 187.34$; $SE = 7.86$) at $F(1,27) = 18.395$, $p < .001$, $\eta^2 = .405$. Neither the *group* main effect, *F*(2,27) = .799, *p* =.46, nor the *group by time* interaction effect was significant, *F*(2,27) = 1.204, *p* = .316.

The Repeated Readings group improved on average 40 words per minute at posttest. Their posttest silent reading rate *(M* = 226.27) is significantly different from their pretest silent reading rate (*M* = 186.26) at *t*(10) = -2.372, *p* < .05, *dz* = 0.72. Similar gains were observed in the Wide Reading group, who on average read 56 more words per minute at posttest (*M*= 249.44) than at pretest (*M*=193.44), with the gain being significant at *t*(8) = -3.142, *p* < .05, *dz* = 1.05. On the other hand, the Vocabulary Study group students whose instruction did not involve any connected text reading, added a non-significant 21 words to their pretest silent reading rate (*M* = 182.3) at posttest (*M* = 203.1), *t*(9) = -1.791, *p* = .107. The mean silent reading rates are displayed in the following figure.

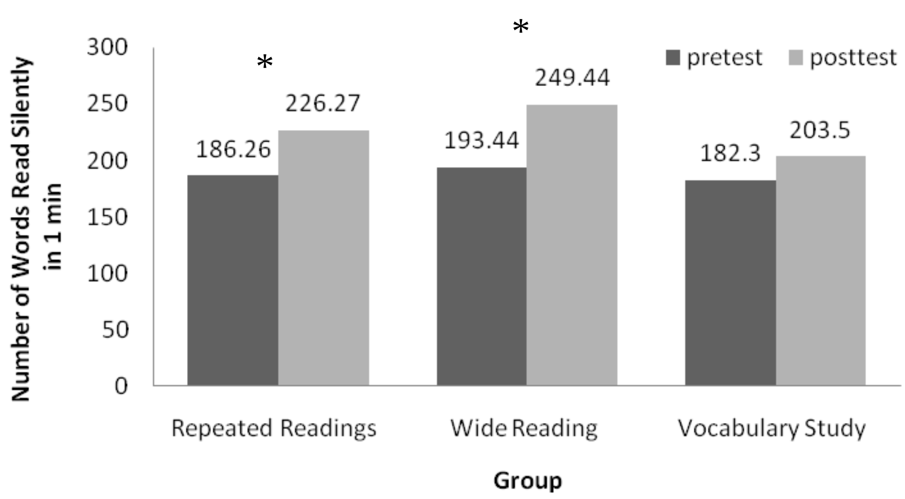


Figure 6. Mean Silent Reading Rate gains by group.

* *p* < .05.

MARSI.

Only the main effect of *time* was significant at $F(1,27) = 6.264$, $p = .019$, $\eta^2 = .188$. Overall time 2 ratings on the MARSI survey were 3.64 ($SE = .103$) compared to 3.38 ($SE = .129$) at time 1. Neither the *group* main effect, $F(2,27) = 1.125$, $p = .339$, nor the *group by time* interaction effect was significant, $F(2,27) = .412$, $p = .666$.

The change for the Repeated Readings group on the MARSI test was only a non-significant .205 points increase at posttest ($M = 3.76$) from pretest ($M = 3.55$), $t(10) = -.939$, $p = .370$. A similar uptick on the MARSI survey ratings was observed for the Wide Reading group, .183 points, from pretest ($M = 3.48$) to posttest ($M = 3.67$), which was non-significant as well at $t(8) = -.948$, $p = .371$. Unlike the fluency training groups, the Vocabulary Study Group reported significantly more metacognitive strategy use at posttest ($M = 3.49$) than at pretest ($M = 3.1$). The .39 difference in response to the items was statistically significant at $t(9) = -4.16$, $p < .005$, $dz = 1.31$. Displayed in the following figure are the pretest-posttest mean comparisons by group.

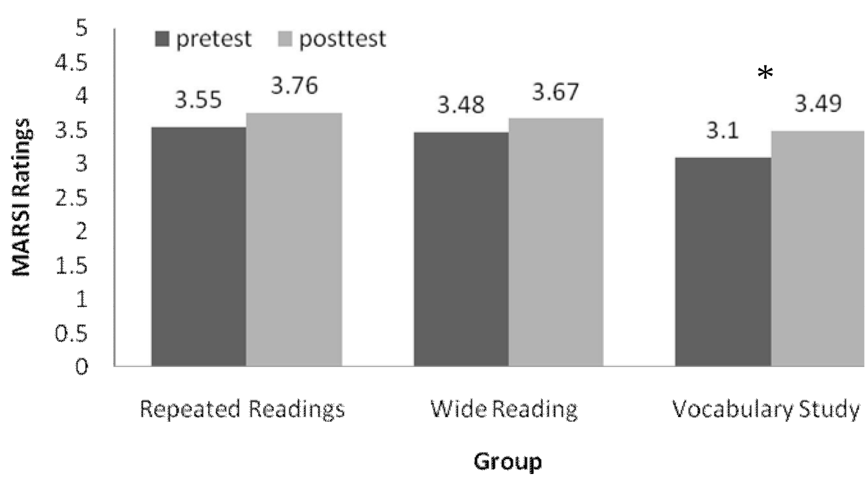


Figure 7. Change in mean MARSI ratings by group.

* $p < .05$.

Maze.

Significant mean differences were observed for the main effect of *time*, $F(1,27) = 29.061$, $p < .001$, $\eta^2 = .518$. Significantly more maze replacements were made across the groups at time 2 ($M = 27.4$; $SE = 1.61$) than at time 1 ($M = 22.1$; $SE = 1.23$). Other effects were not significant: *group* at $F(2,27) = 1.899$, $p = .169$; *group by time* at $F(2,27) = .374$, $p = .691$.

The pretest-posttest difference on the maze test was statistically significant for the Repeated Readings students, $t(10) = -3.46$, $p = .006$, $dz = 1.04$. Repeated Readings participants made 4.27 more maze replacements at posttest ($M = 25.18$) than at pretest ($M = 20.91$). Gains in maze replacements were also observed in the other fluency training group of Wide Reading. Slightly better than the Repeated Readings students, the Wide Reading students made 5.33 more replacements on the maze test at posttest ($M= 31.22$) than at pretest ($M = 25.88$), with this gain being significant at $t(8) = -2.412$, $p < .05$, $dz = 0.8$. Maze gains were not limited to the two fluency training groups; the Vocabulary Study group also achieved gains of similar magnitude (6.3 mazes) on the posttest maze test ($M = 25.8$) relative to the pretest maze test ($M = 19.5$); the gains were of statistical significance at $t(9) = -3.720$, $p =.005$, $dz = 1.18$. Mean gains are displayed in the following figure.

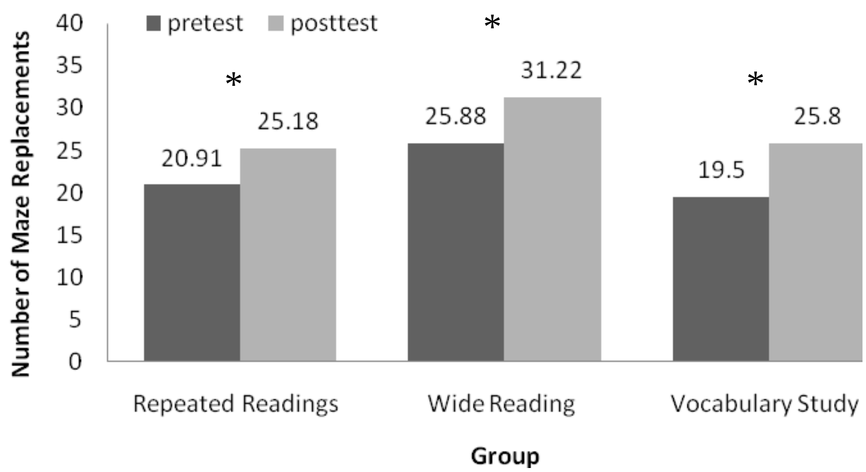


Figure 8. Mean maze gains by group.

* $p < .05$.

RSPAN.

A significant *time* main effect was also the finding on the working memory of RSPAN, $F(1,27) = 6.194$, $p < .05$, $\eta^2 = .187$. A significantly greater number of words were recalled at posttest ($M = 47.84$; $SE = 2.33$) than at pretest ($M = 42.44$; $SE = 2.19$) across the study groups. The main effect of *group*, $F(2,27) = 2.725$, $p = .084$, and the interaction effect of *group by time* were not significant, $F(2,27) = 1.994$, $p = .156$.

On the working memory measure of RSPAN, only the Wide Reading group recalled a significantly larger number of words at posttest ($M = 53.75$) than at pretest ($M = 42.22$), $t(8) = -2.388$, $p < .05$, $dz = 0.795$. Neither Repeated Readings nor Vocabulary Study groups recalled a significantly different number of words at posttest from pretest.

While the Repeated Readings group gained only 1.33 words at posttest ($M_{pretest}$ = 38.11 vs. $M_{posttest}$ = 39.44), the Vocabulary Study group added 3.33 words at posttest ($M_{pretest}$ = 47 vs. $M_{posttest}$ = 50.33). The following figure displays the pretest-posttest comparisons by group.

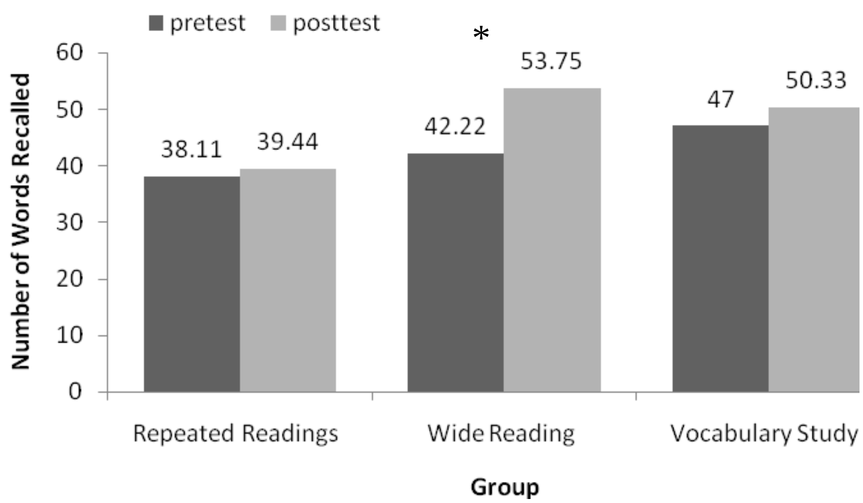


Figure 9. Mean RSPAN gains by group.

* $p < .05$.

TOWRE SWE.

No significant effects were found. No significant overall mean differences were observed (a) from pretest to posttest, $F(1,27) = 1.452$, $p = .239$; (b) across groups, $F(2,27) = 2.342$, $p = .115$, nor were there any differential gains for groups from time 1 to time 2, $F(2,27) = .350$, $p = .708$.

TOWRE PDE.

Similar to TOWRE SWE, no significant *time*, *group*, and *group by time* effects were observed. Overall, groups did not differ in reading aloud lists of non-words within 45 seconds, $F(2,27) = .137$, $p = .872$. Overall performance at time 2 was not significantly different from time 1 performance, $F(1,27) = .622$, $p = .437$. None of the groups performed significantly greater than the others from pretest to posttest, $F(2,27) = .657$, $p = .526$.

Summary and discussion.

Participants as a whole gained at posttest on vocabulary, silent reading rate, maze, and RSPAN. Their MARSI ratings also increased significantly at posttest relative to pretest ratings. Despite these overall gains, no differential gains among groups were observed. In other words, all groups achieved commensurately from pretest to posttest. A summary table of the Repeated Measures ANOVAs lists the significant effect of *time* and non-significant effects of *group* and *group by time* below.

Table 10

(Non-)significant Effects from Repeated Measures ANOVA Results

	Time	Group	Group by Time
NDRC	-	-	-
NDVOC	+	-	-
Silent Reading Rate	+	-	-
MARSI	+	-	-
Maze	+	-	-
RSPAN	+	-	-
TOWRE SWE	-	-	-
TOWRE PDE	-	-	-

Note. NDRC = ND Reading Comprehension Test; NDVOC = ND Vocabulary Test; MARSI = Metacognitive Awareness of Reading Strategies Inventory; RSPAN = Reading Span Test; TOWRE = Test of Word Reading Efficiency Sight Word Efficiency (SWE)/Phonemic Decoding Efficiency (PDE). The plus (+) sign denotes a significant effect whereas the minus (-) sign indicates no significance.

Separate paired-samples *t*-tests were conducted to further explore within-group gains on the DVs (i.e., reading measures), for which a significant *time* effect was detected by Repeated Measures ANOVA. The following table displays the reading measures, on which the groups achieved (non-)significantly from pretest to posttest.

Table 11

Pretest-Posttest Gains by Group

	NDSRR	Maze	NDVOC	RSPAN	MARSI
RR	+	+	-	-	-
WR	+	+	+	+	-
VS	-	+	+	-	+

Note. NDSRR = ND Silent Reading Rate; NDVOC = ND Vocabulary; RSPAN = Reading Span; MARSI = Metacognitive Awareness of Reading Strategies Inventory; RR = Repeated Readings; WR = Wide Reading; VS = Vocabulary Study. The plus (+) sign denotes a significant gain from pretest to posttest; the minus (-) sign denotes a non-significant change.

As is shown in table above, exposure to a wider range of reading materials resulted in significant gains on *more* reading measures than repeated readings of a smaller amount of text or focused study of academic vocabulary. The RR group improved significantly on only Silent Reading Rate and maze. The WR group demonstrated significant increases on all of the measures listed in the table above except for the MARSI survey. For the VS group, significant pretest-posttest changes were observed on Vocabulary, maze and MARSI.

The two fluency programs, RR and WR, were effective in this study in improving poor readers' silent reading rate. The ND Silent Reading Rate outcomes suggest that fluency instruction does lead to reliable gains in the rate at which students read silently in one minute. Both the RR and WR students read at a minimum of 40 more words per minute at posttest. Such significant gains shunned the VS group, whose instruction did not involve reading connected text. It appears, therefore, that a structured fluency program increases speed of poor college readers no matter the design of the fluency instruction; reading repetitively or reading a larger volume of text non-repetitively leads to significant increases in silent reading rate.

Improved reading rate was also the finding of an intervention study comparing the relative effectiveness of repeated to continuous reading on the component skills of second- and fourth-grade readers with and without learning disabilities (O'Connor, White, & Swanson, 2007). In addition to the age difference, O’Connor et al.’s intervention differed from this study in providing corrective feedback during oral reading practice. Both treatment groups of students showed a faster rate of growth on fluency measures than the control group. However, no differential practice (repeated vs. continuous) effects were observed, which is a common finding in the literature (Homan et al., 1993; Rashotte & Torgesen, 1985) amid mixed findings for the superiority of RR over WR (Homan et al., 1993; Kuhn, 2005a; Rashotte & Torgesen, 1985) and findings favoring WR over RR (Kuhn et al., 2006; Mathes & Fuchs, 1993).

All three groups gained significantly on the maze task from pretest to posttest. The maze task has been shown to capture reading processes at both the lower and higher levels all at the same time (see Fuchs et al., 1992; Williams et al., in print) and therefore appears to be sensitive to improvements in reading processes. Silent reading rate gains by the RR and WR groups and vocabulary gains by WR and VS groups may have resulted in increased posttest maze performance observed in all groups. While the RR group’s maze gains may be due to increased reading speed, the VS group’s gains may be due to greater vocabulary knowledge, and the WR group’s maze gains to both increased reading speed and vocabulary knowledge.

Although relatively unsubstantiated, a prevailing conjecture in the literature is that broader exposure to words in varied contexts leads to greater vocabulary acquisition compared to repeated exposure to a smaller amount of text (O'Connor et al., 2007). This conjecture is borne out by the WR group, who made significant gains on vocabulary at posttest. In comparison, the RR condition showed a non-significant change on vocabulary from pretest to posttest. In one of the few studies of the type of reading practice, however, no differences across the fluency groups (i.e., repeated vs. continuous reading) and the control group were found on vocabulary growth (O'Connor et al., 2007).

The significant vocabulary and maze gains observed in the Vocabulary Study students appear to be due to the focused vocabulary study that this group was engaged in. As part of their program, students in this group studied 15 rare academic words and took two quizzes each session. One quiz required them to match 10 of the studied words to their synonyms while the second quiz required 5 words to be matched to their antonyms. Word study and quizzes may have focused the students’ attention in this group on word meaning associations by relating the study words to their synonyms and antonyms. In this respect, the Vocabulary Study program components share construct validity with the ND Vocabulary subtest, which requires students to select the best synonym or the best antonym for a given vocabulary item. Therefore, the Vocabulary Study training seems to have provided targeted practice for participants in this condition and helped raise their sensitivity to word associations that are tested on the vocabulary subtest of the ND Reading Test.

WR is the only group with significant pretest-posttest gains on the RSPAN measure of working memory. Short-term memory (STM) is characterized by limited capacity and rapid decay (Baddeley & Hitch, 1974), which pose limitations to the storage of information during a processing task like reading. In reading, words and phrases are decoded and encoded into phrases and sentences. The intermediate and final products of this process are to be held in memory while the system encodes the propositions. Three strategies are known that minimize the limitations of the short term memory: chunking, rehearsal, and parsing. Breznitz and Share (1992) suggest an additional strategy to circumvent the limitations of short-term memory: accelerating the rate of stimulus presentation. This strategy seems to be achieved only by increased fluency, thereby allowing a greater stretch of the text to be encoded during reading. It appears from the findings of this study that due to exposure to a larger breadth of text, students in the WR group increased their rate of stimulus processing while simultaneously storing the products of intermediate processes. The WR group achieved a significant 11.53-word gain at posttest on the RSPAN from pretest while the other groups' gains were incomparably low: 1.33 words for the Repeated Readings group and 3.33 words for the Vocabulary Study group.

The only group to report a significantly greater observance of reading behaviors on the MARSI survey from pretest to posttest is the Vocabulary Study group. MARSI improvement, however, appears to be an epiphenomenon of the word study that students in this condition completed. It appears that focused attention to word meaning associations induced greater attention to reading comprehension behaviors in the absence of metacognitive instruction. Increased strategy use has mostly been the result of targeted instruction in which a number of select strategies are modeled and practiced (Baker & Brown, 1984; Bereiter & Bird, 1985; Palincsar & Brown, 1984; Rosenshine & Meister, 1994); among most notable are Palincsar and Brown's research which combined four major strategies (e.g., summarizing, clarifying, questioning, predicting) in an instructional protocol called Reciprocal Teaching and Bereiter and Bird's research which resulted in the identification and teaching of effective strategies used by accomplished readers. It is a rare finding that students' awareness and use of reading strategies are increased by vocabulary instruction.

Why were there no significant effects on the ND Reading Comprehension subtest? The reading comprehension performance of Repeated Readings students declined by 2.55 questions at posttest (M = 28.73) from pretest (M = 31.27); for the Wide Reading group there was a .44 questions increase at posttest (M = 34.44) from pretest (M = 34); and the Vocabulary Study group's reading comprehension dropped to an average of 29 questions correct at posttest from 36 questions correct at pretest. Lack of differential gains in reading comprehension is a common finding in studies comparing repeated readings to wide reading (Homan et al., 1993; Kuhn et al., 2006; Mathes, 1993; O'Connor et al., 2007; Rashotte & Torgesen, 1985), with the general conclusion being that reading a greater volume of text non-repetitively does not necessarily lead to any larger gains in the reading comprehension of struggling readers than repeated readings of a smaller amount of text. The only dissenting finding to date has been reported by Kuhn (2005). Second-graders, who participated in either a wide reading condition or a variation of the repeated readings condition, read aloud passages from the Qualitative Reading Inventory (QRI, 1988) and Qualitative Reading Inventory II (QRI-II, 1995) and answered comprehension questions about the passages as part of the assessment of fluency program effectiveness. Although both groups were rated more fluent in their oral reading than students in a listening-only condition and control students, improved comprehension was achieved only by the wide reading group in response to questions about the QRI and QRI-II passages.

Previous research adopted the notion that due to broader exposure to unique words used in different contexts greater gains in reading comprehension would be obtained from the WR condition than the RR condition (Kuhn, 2005a; O'Connor et al., 2007). However, none of the groups improved significantly from pretest in this study. These findings provide further evidence that reading comprehension is a multi-componential skill (Baddeley et al., 1985; Guthrie, 1973; Palmer, Mcleod, Hunt, & Davidson, 1985), which may require that all components be targeted for gains on a standardized test like the ND Reading Comprehension Test to materialize. Simply targeting the fluency component may not lead to gains sufficient enough to ameliorate deficits in other components including vocabulary knowledge. For example, despite sizeable RSPAN gains, the WR group did not add significant gains to their reading comprehension at posttest although parallel gains in reading comprehension were an expected outcome for this condition due to the strong working memory-reading comprehension relationship (Daneman & Carpenter, 1980). Daneman and Carpenter found the listening span scores and the Verbal SAT scores in a group of college readers to correlate .53.

Lack of WR gains in reading comprehension in spite of the significant working memory gains is also further evidence that fluency training effects may be more tangible from tasks that tap underlying processes of reading comprehension than from global measures of standardized reading comprehension tests. In support of this argument, Daneman and Carpenter report stronger correlations between working memory and more specific comprehension tests of answering fact questions (r = .67), answering pronoun reference questions (r = .72), and abstracting a theme from the spoken narrative passage by providing a title (r = .82).

However, there are signs of *collateral* gains in the reading components of the WR group, whose fluency instruction involved reading four times as many words as the RR group. In addition to improvements in silent reading rate and maze, also observed in the RR group, the WR students ended the training with greater vocabulary knowledge and improved working memory efficiency. The WR gains in multiple reading measures portend increased achievement on global reading measures, such as the reading comprehension subtest of the ND Reading Test, in longer-duration interventions.

Why did overall gains eschew the TOWRE subtests? None of the groups' pretest and posttest scores differed to a significant degree on the TOWRE SWE measure of timed sight word recognition. The RR group lost 2.6 sight words at posttest (M = 78.8) compared to the pretest (M = 81.4), $t(10)$ = .905, p = .387. There was *zero* change in the WR group's posttest achievement on this measure from pretest (M = 81.33), $t(8)$ = 0. The performance of the VS group, on the other hand, declined 2.33 words at posttest (M = 72.33) from pretest (M = 74.66), $t(9)$ = 1.317, p = .220. Similar to the results of TOWRE SWE, no significant changes were detected for all three groups from pretest to posttest on TOWRE PDE. The RR group read aloud 3.4 fewer nonwords correctly in 45 seconds at posttest (M = 24) than at pretest (M = 27.4), $t(10)$ = 1.475, p =.171. The WR group read aloud .22 more non-words at posttest (*M* = 28.11) than at pretest (*M* = 27.89), $t(8) = -.116$, $p = .910$. Moreover, the VS group declined on posttest (25.55 vs. 25.77) by .22 non-words, $t(9) = .074$, $p = .943$.

These results add to the mixed findings of fluency effects on word recognition reported by recent research studies. While the results from this study are discrepant with Kuhn's (2004) findings of improved recognition of words in isolation (measured by TOWRE) by both wide-reading and a variation of repeated readings relative to a silent reading and control groups, they are in agreement with O'Connor et al.'s (2007) findings of no differences between the wide and continuous reading groups on the Word Identification subtest of the Woodcock Reading Mastery Test-NU (WRMT-NU; Woodcock, 1998). Similar to O'Connor et al., Homan, Klesius, and Hite (1993) found no gains in word recognition across the groups. Equal gains in accuracy and speed were detected for a repeated readings group and an assisted, non-repetitive oral reading strategies (i.e., echo reading, unison reading, and cloze reading) group.

Lack of feedback may be a reason that students did not improve on their efficiency of reading words and non-words in isolation. Because the fluency training was conducted silently, students did not receive corrective feedback on hard-to-decode or unfamiliar words. Providing corrective feedback and engaging students in reading aloud were two recommendations made by several reviews of fluency-building programs (Kuhn & Stahl, 2003; Mercer & Campbell, 1998; Meyer & Felton, 1999; Wolf & Bowers, 1999), which were not implemented in this study due to logistical constraints. Improvements in single (non-)word reading have been reported by studies which incorporated these recommendations, i.e., corrective feedback and teacher support in the form or choral and echo reading (see Kuhn, 2005a).

Relatedly, during silent reading students may choose to skip over hard-to-decode words and still be able to derive a correct meaning for such words from the context of the sentence. This was among the findings of a study by Juel and Holmes (1981), who found faster reading latencies on sentences read silently than on sentences read aloud. It appears from their findings that subjects seemed to “skip over” words that they found hard to decode in silently read sentences as opposed to those they encountered in sentences read aloud. Despite this latency difference across the two modalities, commensurate comprehension in both conditions was achieved. Therefore, it is most likely that silent reading does not lend to focused attention to decoding words that is strictly induced in reading aloud and thus leaves students’ decoding errors uncorrected and their decoding skills unpracticed. By bypassing the decoding of difficult words, students may choose instead to commit an unfamiliar word to their sight vocabulary. Although they may be able to derive the meaning of a difficult word from the sentence context, the lexical representation they construct for the word lacks the phonological specificity that is necessary for the quality of the lexical representation. According to Perfetti’s (2007) theory of lexical quality, word knowledge is comprised of specificity in orthographic, phonological, and semantic representations. Lacking quality and specificity in one of the three representations, the word cannot be said to be unambiguously known.

Research Question C: How Do the Study Groups Compare at Posttest on Reading Ability Measures?

To compare the groups on posttest measures, Analysis of Covariance (ANCOVA) was utilized to reduce systematic bias that may emanate from pretest performance. The purpose of using ANCOVA was to control for groups' differences on pretest measures and thus avoid the confounding of the treatment effects with initial differences. Without this statistical control, it would not be possible to differentiate between the fluency training and the initial differences (at pretest) as the source of differences observed at posttest. If not controlled for, initial differences may carry over to posttest performance (Stevens, 1999). Separate ANCOVAs were computed for each dependent variable measured at posttest with data from its respective pretest measure serving as the covariate. No significant differences were detected in the adjusted group means on each of the ANCOVA computations: Reading Comprehension, $F(2,26) = 1.135$, $p = .337$; Vocabulary, $F(2,26) = .178$, $p = .838$; Silent Reading Rate, $F(2,26) = 1.149$, $p = .333$; *Maze*, $F(2,26) = .375$, $p = .691$; MARSI, $F(2,26) = .044$, $p = .957$; RSPAN, $F(2,26) = 2.908$, $p = .073$; TOWRE SWE, $F(2,26) = .806$, $p = .457$; TOWRE PDE, $F(2,26) = .662$, $p = .524$. In other words, none of the study groups with pre- and post-test data (i.e., RR , WR, and VS) demonstrated an advantage over the others in posttest achievement once their pretest scores were accounted for through the ANCOVA statistical procedure.

Research Question D: Does Fluency Training Raise Sensitivity to Textual Inconsistencies? Which Fluency Program (RR vs. WR) Leads to Greater Gains?

The answer to this question requires a comparison of the pattern of reading times from pretest to posttest. First, reading time data from pretest will be tabulated and analyzed for significant effects before the pattern of results are compared to those from posttest.

Groups' reading times from the pretest Error Detection Task are depicted in the table below. A 4 (group) X 2 (consistency) X 2 (coherence) Repeated Measures Analysis of Variance (ANOVA) on target-action reading times was conducted. Group (Repeated Readings vs. Wide Reading vs. Vocabulary Study vs. Comparison Group) was a between-subjects factor, and consistency (consistent vs. inconsistent), and coherence (global vs. local) were within-subjects factors. All latencies larger than 6,500 ms and shorter than 200 ms were treated as missing data and excluded from the analyses.

Table 12

Reading Time Latencies by Group at Pretest

	Local			Global		
	Consistent		Inconsistent	Consistent		Inconsistent
RR (*n*=11)	2168.06 (699.64)	<	2358.73 (839.85)	1991.95 (826.06)	<	2148.47 (862.77)
WR (*n*=9)	2154.74 (565.62)	>	2120.69 (798.16)	2058.84 (451.94)	<	2291.54 (910.12)
VS (*n*=10)	2292.74 (469.53)	<	2432.92 (567.19)	2368.56 (521.98)	<	2554.11 (809.25)
CG (*n*=13)	1338.86 (652.05)	<	1435.42 (558.22)	1307.23 (520.43)	<	1393.70 (555.67)

Note. RR = Repeated Readings; WR = Wide Reading; CG= Comparison Group. Standard deviations are reported in parentheses.

The analysis revealed reliable main effects of *consistency*, $F(1,39) = 5.960$, $p < .05$, partial eta squared = .133, and *group*, $F(3,39) = 6.230$, $p = .001$, partial eta squared = .324. No other main or interaction effects were significant. The *consistency* effect was due to overall slower reading times on inconsistent sentences ($M = 2092$; $SE = 109.83$) than consistent sentences ($M = 1960$; $SE = 90.103$), and the *group* effect was due to the significantly faster response times by the Comparison Group ($M = 1369$; $SE = 174.36$) relative to Repeated Readings ($M = 2167$; $SE= 189.57$), Wide Reading ($M = 2156$; $SE= 209.57$) and Vocabulary Study ($M = 2412$; $SE= 198.82$) groups.

Although non-significant, Table 12 depicts a pattern of reading times in which all groups experienced comprehension difficulty reading the target sentence in the global-inconsistent condition, indicated by larger reading times, compared to the global-consistent condition. The same pattern is observed in the local condition, except for the

WR group, who on average spent more time on the locally consistent passages than locally inconsistent passages. While the pattern of reading times in the local condition, with the exception of the Wide Reading group, complies with previous research findings that poor readers exhibit an inconsistency effect in the local condition, it was in opposition to the predictions of previous research that they do so in the global condition. Equivalent reading times have been suggested for less skilled readers on global-consistent and global-inconsistent sentences (Long & Chong, 2001).

In the Long and Chong (2001) study, which was a mixed factorial design with the reader group (good vs. poor comprehenders) as the only between-subject variable and with character (first vs. second character), distance (global vs. local), and consistency (consistent vs. inconsistent) as the within-group variables, the following pattern of reading time data was observed as listed in Table 13.

Table 13

Pattern of Reading-Time Data on the Error Detection Task in Long and Chong (2001)

	First Character (Ken)						Second Character (Mike)					
	Global			Local			Global			Local		
	Con		Incon	Con		Incon	Con		Incon	Con		Incon
GR	Y	<	X	Y	<	X	Y	=	X	Y	=	X
PR	Y	=	X	Y	<	X	Y	=	X	Y	<	X

Note. GR = Good readers; PR = Poor readers; Con = Consistent; Incon = Inconsistent. The letters X and Y denote reading latencies.

Long and Chong (2001) found good comprehenders to be slower in reading the target sentence when it was inconsistent with the first-character description. Good comprehenders demonstrated slower reading on the target sentence both when the target sentence and the appropriate character (first character Ken) elaboration were separated out by a number of filler sentences (i.e., the global condition) and when they were separated out only by one sentence. Poor comprehenders seemed to detect the error (inconsistency) in the local condition just like the good readers; however, they failed to do so in the global condition where the elaboration and the target sentence which introduced the error were separated by a longer section of filler sentences.

For the second character, which was irrelevant to the target sentence, the groups differed only in the local condition. While the good comprehenders did not show an inconsistency effect for the second character, the poor readers did read the inconsistent sentence slower for the second character in the local condition. We know from the first-character data analysis that the poor readers' failure to detect the error may have stemmed from an inability to reactivate the relevant information that was eliminated from working memory by a filler section of sentences. This conclusion can be used to explain why they acted like the good readers in the global condition. The fact that the poor readers were slower to read the target sentence in the local condition for the second character is an indication that their mental models are not well specified for the correct character; the second character is irrelevant for the action in the target sentence, which is performed by the first character.

Long and Chong's (2001) second character condition is irrelevant to the current study; reading time data from that condition was not considered in the interpretation of the data from the current study, which presented passages with only one elaborated character who executed the action described by the target sentence. The reading time data from the current study were compared to those obtained by Long and Chong on the first-character passages. Unlike Long and Chong's students, all poor comprehenders (RR, WR, VS students) in this study displayed a non-significant inconsistency effect in the global condition. In the local condition, though non-significant as well, the expected pattern was broken by the Wide Reading group, whose reading times were inexplicably larger on the consistent than the inconsistent target sentences. It was expected from Long and Chong findings that all subjects display an inconsistency effect in the local condition.

How did the training affect the pattern of results obtained from the pretest measure of the Error Detection task? The following table shows the expected changes in the pattern of reading times as a result of the training that subjects in the Repeated Readings and Wide Reading conditions experienced. Recall that the Vocabulary Study group showed an inconsistency effect in both the local and global conditions, albeit non-significant. Although the inconsistency effect in the local condition was an expected outcome, it was contrary to expectations that this group of poor readers show an inconsistency effect in the global condition. Because the Vocabulary Study Group's training did not involve reading connected text, their pattern of reading time data from posttest should mirror those from pretest.

Table 14

Hypothesized Changes in the Pattern of Reaction Time Data on the Post-test Error Detection Task

	Global			Local		
	Consis		Inconsis	Consis		Inconsis
RR	Y	<	X	Y	<	X
WR	Y	<	X	Y	<	X
VS	Y	<	X	Y	<	X
CG PRETEST[+]	Y	<	X	Y	<	X

*Note.*The letters X and Y denote reading latencies. Consis = Consistent; Inconsis = Inconsistent.

+ Comparison Group did not take the posttest ERDE task. Comparison Group's pretest ERDE data are included in this table for comparison purposes. Comparison Group's pretest ERDE data were not included any analyses of treatment effects.

The pattern of reading times obtained on the posttest ERDE is depicted below in Table 15.

Table 15

Reading Latencies by Group on Posttest Error Detection Task

	Local			Global		
	Consistent		Inconsistent	Consistent		Inconsistent
RR (*n*=11)	2096.4 (772.248)	>	2055.17 (728.693)	1962.93 (747.272)	<	2026.08 (843.171)
WR (*n*=9)	1647.1 (437.399)	<	1990.918 (501.318)	1833.71 (659.926)	>	2038.65 (605.028)
VS (*n*=9)	2314.95 (722.497)	>	2170.69 (680.603)	2144.26 (592.958)	<	2398.95 (798.582)
CG PRETEST [+]	1338.86 (652.045)	<	1435.42 (558.223)	1307.23 (520.425)	<	1393.70 (555.668)

Note. Standard deviations are reported in parentheses.

[+] Comparison Group did not take the posttest ERDE task. Comparison Group's pretest ERDE data are included in this table for comparison purposes only. Comparison Group's pretest ERDE data were not included any analyses of treatment effects.

A 3 (group) X 2 (consistency) X 2 (coherence) X 2(time) Analysis of Variance (ANOVA) with repeated measures was conducted on target-action reading times. Group (Repeated Readings vs. Wide Reading vs. Vocabulary Study) was a between-subjects factor, and consistency (consistent vs. inconsistent), coherence (global vs. local), and time (pretest vs. posttest) were within-subjects factors. All latencies larger than 6,500 ms and shorter than 200 ms were treated as missing data and excluded from the analyses.

The analysis revealed reliable main effects of *consistency*, $F(1,27) = 11.573$, $p <.005$, partial eta squared = .300, and *time*, $F(1,27) = 4.497$, $p < .05$, partial eta squared= .143, and a reliable interaction effect between *coherence* and *group* at $F(2,27)= 3.767$, $p< .05$, partial eta squared = .218. No other main or interaction effects were significant. Lack of significant three- and four-way interactions render the reading time data tabulated in

Table 14 unreliable due to sampling error. In other words, the data listed may be chance occurrence and not represent the true reading times that struggling college readers who participated in this study may ordinarily exhibit at the end of a fluency training program. The group means are depicted only to illustrate the pattern of response times from the groups.

The *consistency* effect was due to overall slower reading times on inconsistent (*M*= 2216; *SE*= 121.33) than consistent (*M* = 2086; *SE* = 105.13) sentences, and the *time* effect was due to the significantly faster overall reading times at posttest (*M* = 2057; *SE* = 119.34) than at pretest (*M* = 2245; *SE*= 121.48). Faster reading times in the global condition (*M* = 2032; *SE* = 195.63) compared to the local condition (*M*= 2170; *SE*= 177.21) for Repeated Readings students; in the local (*M*= 1978; *SE*= 195.91) than global condition (*M* = 2056; *SE* = 216.28) for Wide Reading students; and in the local (*M*= 2303; *SE*= 185.86) than global condition (*M*= 2366; *SE*= 205.18) for Vocabulary Study students led to the significant *group by coherence* interaction.

In sum, the pretest and posttest ERDE administrations resulted in non-significant, *anomalous* patterns of reading time for all groups. The anomaly observed in the patterns of reading times is attributable to two main confounds. First, limited English proficiency of some of the participants speaking English as a second language may have undermined their ability to read for meaning. Unlike this study, previous research investigations of the inconsistency paradigm used only native speakers of English as participants (Albrecht & O'Brien, 1993; Cook et al., 1998; Hakala & O’Brien, 1995; Long & Chong, 2001; O'Brien & Albrecht, 1992; O'Brien et al., 1998). Second, although administered individually, the ERDE task was taken by participants in a tutoring room, which was used to accommodate eight to 10 participants taking the ERDE task and other computer-based tests (i.e., TOWRE, RSPAN). Students taking the ERDE task may have been distracted by the other participants taking the TOWRE subtests, which required reading aloud (non)words to a microphone or students taking the RSPAN task, for which participants recalled words from trials to a microphone.

CHAPTER SIX

CONCLUSION

This study investigated the effects of fluency training on less skilled college readers' ability to integrate information in narrative passages in an error detection task in which textual inconsistencies are used to induce comprehension monitoring (Albrecht & Myers, 1998; Albrecht & O'Brien, 1993; Garner, 1981; Long & Chong, 2001; O'Brien & Albrecht, 1992; O'Brien et al., 1998; Oakhill et al., 2005). The task involves collecting reading time data from subjects who read short passages on a computer screen line by line at their normal pace with comprehension—a comprehension question that appears at the end of the passage ensures subjects' attention to the story. Subjects use a line-advance key to retrieve the lines of the passage while their reading times on each line are recorded.

In order to observe metacognitive behavior (longer reading times on certain lines compared to others), experimental passages are manipulated such that two sentences in the passage are made to contradict each other. This manipulation reliably redflags good and older comprehenders (compared to younger and poor comprehenders) and leads to them to evaluate the state of their comprehension by spending relatively longer time on the target sentence, which causes the inconsistency (August et al., 1984; Baker, 1985; Baker & Anderson, 1982; Wagoner, 1983); the elapse of time spent reading the target sentence is taken as an indication of subjects' cognitive processes to recover from the comprehension break.

Manipulations of the distance between the inconsistent segments in an error detection task have revealed further insights into the extent to which subjects are able to construct coherent mental models. This has resulted in a distance effect, which robustly differentiates skilled from less skilled readers (Long & Chong, 2001). While both groups of readers ably monitor their comprehension when the inconsistent sentences are adjacent, only skilled readers are able to display evidence of comprehension monitoring at the global condition where the inconsistent sentences are separated out by a filler section of intervening sentences.

Two fluency interventions were implemented to help poor readers "develop" the inconsistency effect in the global condition as a result of increased processing efficiency. Two approaches that provide practice with print, a Repeated Readings program and a Wide Reading program, were compared in struggling college readers enrolled in a Reading Course at a community college in the Southeast. In the Repeated Readings condition students read a grade-level passage four times back to back and answered 10 comprehension questions about the passage per session. In the Wide Reading condition, however, four different passages were read per session each once. For each passage three comprehension questions were answered by the WR students. Reading times per reading were recorded as well as comprehension accuracy score per session. Because the study was conducted in students' class time, a Vocabulary Study condition was designed for students randomly assigned to serve as the control condition. The VS students' training did not involve reading connected text; students in this condition studied lists of 15 words per session, took two quizzes (a synonyms quiz and an antonyms quiz), and completed a word card for unfamiliar words.

Data were analyzed for general and specific training effects on reading rate, reading comprehension, working memory and error detection. In light of theoretical insights (Daneman, 1987; Laberge & Samuels, 1974; Perfetti, 1985) and recent research findings (Kuhn, 2005b; Kuhn et al., 2006; Kuhn & Stahl, 2003), it was expected that practice provided by extensive reading of greater amount of text may lead to similar or larger gains in word recognition, text reading fluency, vocabulary knowledge, and reading comprehension than repeated readings of a smaller amount of text. Due to processing efficiency gains, struggling readers were expected to allocate adequate attentional resources to monitor their comprehension and construct coherent mental models by reliably detecting errors on the error detection task.

The analyses revealed (a) non-significant training effects on reading comprehension and word recognition, (b) significant within-group gains for the Wide Reading condition in vocabulary, working memory, silent reading rate, and maze, (c) significant within-group gains for the RR students in silent reading rate and maze, (d) significant increases on maze, vocabulary and MARSI for the VS group at posttest, and (e) non-significant, *anomalous* reading time patterns across groups on the ERDE task both at pretest and posttest.

On pretest ERDE, all participants were expected to display a pattern of responding (i.e., RR, WR, VS) obtained from a group of college students defined as poor comprehenders by Long and Chong (2001). Although non-significant, *anomalous* patterns of reading time data were observed on the pretest ERDE task for all poor readers.

Except for the WR group, all groups demonstrated a non-significant inconsistency effect in the global and local conditions with longer reading times spent on inconsistent sentences compared to consistent sentences. The WR group's reading times were larger on the local-consistent sentences than their reading times on the local-inconsistent sentences; their reading times in the global condition conformed to those of other groups. These findings (although non-significant) are contradictory to the reading-time pattern predicted by prior research, which has found an inconsistency effect in the local condition but not in the global condition for poor readers (Albrecht & O'Brien, 1993; Hakala & O'Brien, 1995; Long & Chong, 2001).

It was hypothesized that by means of the fluency training, the anomalous pretest reading time patterns on the ERDE task would normalize to skilled comprehenders' pattern of responding as a result of increased processing efficiency at post-test. Operationally, longer reading latencies were expected by RR and WR students on local and global inconsistent sentences. Normalization in reading patterns was expected for fluency participants whose increased fluency may provide them with greater working memory capacity, thus enabling them to monitor their comprehension for comprehesion breaks with greater facility. This reading-time normalization was not expected for the Vocabulary Study group, whose instruction did not involve connected text reading. Although non-significant, the posttest reading time data from the ERDE task does not show any indication of developing an inconsistency effect both in the local and global conditions across all participants.

Despite the null findings on the ERDE task, the reading gains made by the WR group in only nine sessions are encouraging. The WR group made gains in vocabulary knowledge and on tasks that require processing efficiency (i.e., working memory, silent reading rate, and maze). The increased processing efficiency of the WR group was expected to extend to the ERDE task, in which increased working memory capacity is assumed to underlie the ability to detect a comprehension break caused by a target action inconsistent with the earlier-presented descriptions of the character. The WR group's increased processing efficiency did not result in reliable integrative skills (on the ERDE task) that have been shown to distinguish the pattern of responding by skilled readers from that of less skilled readers.

Improvements in integrative processes of comprehension may be more reliably achieved at the local levels of proposition construction on a more sensitive measure such as the maze test. The maze was the only measure to pick up training effects in this study; both RR and WR groups made significantly more maze replacements at posttest than at pretest. Equivalent maze gains were also observed in the VS condition. It is very likely that studying word meanings through semantic enrichment induced greater integrative processes in VS students. In doing the synonym and antonym quizzes, students may have developed a greater tendency for semantic associations between words. This increased sensitivity may have led them to execute better integrative processes that characterize the maze task. In other words, local level activation and suppression skills of students may have benefitted from the focus that vocabulary study placed on meaning enhancing associations during vocabulary study.

The results reported in this study, however, may have been affected by a number of limitations. First, a major limitation concerns using a pre- and post-test design to examine changes in reaction time data derived from the error detection task. Reaction time data have been rarely analyzed to reflect changes over time in the literature. However, using a comparison group of successful readers, as was done in this study, may serve to provide a norm-like data source to increase the utility of reaction time data in a pre- and post-test design.

Second, in prior investigations of the inconsistency paradigm, data were collected from native speakers of English. Due to logistical constraints, this study recruited a mixed sample of native and nonnative speakers of English enrolled in a community college in Southeastern US. The “anomalous” reading time data observed on the Error Detection task may have stemmed from the non-native speaking students’ unfamiliarity with the topics covered in the ERDE passages and lack of experience reading English narratives.

Third, although nine sessions of fluency training were found to be adequate for training effects to stabilize, the short duration of the fluency training may not have been adequate for training effects to consolidate. A longer training program may see more stable outcomes in participants’ performance on the psycholinguistic tasks from which reading latencies are gathered, such as the ERDE task used in this study.

Finally, psycholinguistic data are conventionally collected in sound-insulated quiet rooms where distraction due to noise is minimized. Reducing distraction is necessary to collecting reliable reading time data from such tasks as the Error Detection paradigm. Due to logistical constraints, the ERDE task in this study was taken by the participants in rooms that were made to accommodate about eight to 10 participants. Because the order of the computerized tasks was counterbalanced, participants taking the

ERDE task may have been distracted by other participants taking the RSPAN task or the TOWRE subtests, for which oral responses are recorded.

Implications

Instruction

Fluent reading ability is as much an outcome of reading practice as it is an antecedent (Pinnell et al., 1995). The more access students have to reading opportunities and the more time they spend reading, the more fluent they become (Allington, 1984; Biemiller, 1977-1978; Cunningham & Stanovich, 1991, 1997; Krashen, 1993; Nagy & Anderson, 1984; Stanovich, 1986). By reading avidly, students are exposed to a wide gamut of language structures, language in use and syntactic elements that they may not encounter in spoken language (R. C. Anderson, Wilson, & Fielding, 1988; Hayes & Ahrens, 1988). Fluent readers choose to read because of the reinforcing past experiences they have had. By reading more, fluent readers enhance their reading skills and are more likely to engage in further reading. Disfluent readers, on the other hand, avoid reading and prefer activities that demand less effort (Daly, Chafouleas, & Skinner, 2004; Nathan & Stanovich, 1991).

As has been shown in prior research, poor reading ability can be effectively improved through fluency programs, which result in improved reading rate, prosodic reading, and comprehension as well as greater desire to read (Dowhower, 1987; Herman, 1985; National Institute of Child Health and Human Development, 2000a), with the effects being larger from reading connected text than practicing lists of words (Fleisher et al., 1979). Furthermore, wide reading may prove more advantageous to reading ability than repeated readings (Kuhn, 2005a; Kuhn et al., 2006; Kuhn et al., 2000).

Substantiating recent findings regarding the salutary effects of exposure to a wider range of reading materials, The WR condition in this study resulted in significant gains on more reading measures than repeated readings of a smaller amount of text or focused study of academic vocabulary. The WR students on average (a) added 5.55 words to their pretest vocabulary score at posttest, (b) increased their silent reading rate by 56 words at posttest, (c) made 5.33 more replacements on the posttest maze test, and (d) recalled 11.53 more words on the working memory measure of RSPAN at posttest. By contrast, significant gains achieved by the RR group are limited to two reading measures: a 40-wpm gain on the ND Silent Reading subtest and a 4.27-maze gain on the maze test. On the other hand, the VS group, who only studied college-level vocabulary, gained 6.6 words on the ND Vocabulary subtest and 6.3 mazes on the maze test at posttest. Moreover, the VS group is the only group to report significantly more metacognitive strategy use at posttest.

In light of these findings, a Wide Reading fluency program is suggested for college readers who do poorly on standardized reading comprehension tests and show signs of limited skills in reading fluency. A Wide Reading program may be implemented with individual students or a group of students in the vicinity of 20 minutes. Binders may be compiled for individual participants consisting of a progress chart, a reading time-rate conversion table, instructional-level passages, and the answer key. Perennial issues of classroom management may cease to be a concern for teachers with increased familiarity with the program routine following the first few sessions.

Furthermore, the gains in reading skills may be augmented by combining the WR instruction with focused vocabulary study. In addition to being the only condition to show greater awareness and use of meta-cognitive reading strategies at posttest, the VS condition resulted in vocabulary and maze gains. Such a combination of wide reading and vocabulary study may prove even more powerful in increasing struggling college readers' reading skills and verbal ability.

Research

On the basis of the findings from this study, it is recommended that lower level reading processes be targeted to discern training effects. At the higher levels, i.e., discourse processing, the effects of training may be harder to isolate due to myriad sources of variance. At the higher grades, language comprehension and vocabulary knowledge play a more significant role in determining reading achievement (Gough et al., 1996). A more appropriate candidate for future investigations may be the processes that occur during parsing and propositional encoding. It was suggested by the pioneers (Dahl, 1974) of Repeated Readings and implied subsequently (Schreiber, 1980) that fluency increases improve the lower level processes of parsing, i.e., proposition construction and integration. Extant psycholinguistic tasks should be employed to identify the improvements to local level processing that are due to fluency training.

For example, the Till, Mross, and Kintsch (1988) paradigm could be used to examine training effects in less skilled readers. In the Till et al. study, students completed a lexical decision task following the presentation of a sentence that ended with a prime. Among the targets in the lexical decision task were associates of the prime word (the sentence final word) and a topical inference word for the sentence. The integrative process of spreading activation resulted in the deactivation of inappropriate associates after a delay of 350ms following the sentence offset. The delay in making a topical inference took subjects 150ms longer than the selection of the appropriate associate. That is, subjects required as long as 500 ms to deactivate an unrelated topical inference word relative to a related inference word. The Till et al. paradigm could be used to examine the patterns of deactivations of inappropriate associates and unrelated topical inference words in less skilled readers as a function of fluency training. Similar to the current study, a group of skilled readers could be recruited to provide comparison data.

Local level parsing processes require integrative skills that may be executed slower (Long et al., 1997), “sluggishly” (Perfetti, 2007) or “shallowly” (Hannon & Daneman, 2004) by less skilled readers. Using the Till et al. paradigm, Long and colleagues had skilled and less skilled readers study 2-sentence passages and press YES if a test word appeared in the passage or press NO if it did not. Appropriate sense of the ambiguous homograph (e.g., *mint* with two meanings, see Table 16 below) was responded to faster by both skilled and less skilled readers than the inappropriate sense; no significant differences were noted between the groups in response times.

Table 16

Example Passages and Target Items from Long et al. (1994)

Passage	Target items	
	Prime	*Topic*
The townspeople were amazed to find that all the buildings had collapsed except the *mint.* Obviously, it had been built to withstand natural disasters.	money	earthquake
Thinking of the amount of garlic in his dinner, the guest asked for a *mint.* He soon felt more comfortable socializing with the others.	candy	breath

However, reliable differences were observed in the topic condition: Appropriate topic words took longer to reject and caused significantly more errors for the skilled group than the less skilled group. No response time differences were observed for the less skilled readers between appropriate and inappropriate topic words in both reading times and error rates. Plausibly, the skilled readers took longer and made more errors in rejecting topic words (that did not appear in the context of the sentences) because these words were incorporated into the topic-related inferences they made while reading the sentences. Such group differences, however, disappeared when skilled and less skilled readers were tested on their knowledge of the passages offline; both groups made correct topic-related inferences (Long et al., 1994; Long et al., 1997). It appears therefore that less skilled readers are delayed in local level integrative processes, the processes that are deployed to reinforce activated memory nodes that are appropriate and deactivate those that are inappropriate.

Alternatively, integrative processes could be investigated across skilled and less skilled readers in a meaning fit judgment task, which measures subjects' susceptibility to interference due to associates of a homograph (Gernsbacher et al., 1990). In this task, subjects verify if a test word following a sentence match the meaning/context of the sentence. A NO response is expected for both of the following trials because the test word *ace* does not fit the meaning of either sentence. While *ace* is an associate of *spade*, it is not semantically associated with *shovel*.

11a. He dug with spade.

ace

11b. He dug with shovel.

ace

Gernsbacher and colleagues (1990) found both skilled and less skilled readers to experience significant interference immediately after reading the test word *ace;* inappropriate meanings were still highly active at the immediate interval. At the one-second delay, however, the interference was only experienced by the less skilled readers, for whom the inappropriate meaning continued to remain activated. In other words, skilled readers were able to suppress the inappropriate meaning at the delay, but not the less skilled readers. In fact, less skilled readers were still experiencing a significant amount of interference after the delay.

The results indicate that sentential integrative processes took the less skilled readers longer to complete. While the skilled group finished this process after the immediate interval (200 ms), the less skilled group was still engaged in the process at least at the delay condition, about a second later. Because both groups were still processing the sentence at the immediate interval, they were vulnerable to the interference from an associate word. However, the associate ceased to interfere for the skilled readers at the delay who had by this time finished processing the sentence and freed their working memory. Less skilled readers, on the other hand, never finished constructing a coherent sentence representation, even until the delay, and therefore were still vulnerable to the interference caused by the associate *ace* of the sentence-final word *spade*.

The effects of fluency training, which has been claimed to expedite the local level (sentential) integrative processes, could be tangibly observed with psycholinguistic paradigms that induce inference making (Gernsbacher et al., 1990; Long et al., 1994; Long et al., 1997; Long et al., 1999). As well, the maze test, which has been found to be a sensitive measure of local (and global) reading processes (Fuchs & Fuchs, 1992; Williams et al., in print), could prove instrumental in this endeavor. The maze test should be compared to the psycholinguistic measures of sentence processing in which proposition construction and integration processes are sufficiently isolated. Training effects may be more evident on measures that are sensitive to local level processing than on measures of discourse processes, the measurement of which are muddled by unexplained sources of variance.

Attesting to the potential utility of the maze task for this purpose are significant correlations between the maze test and multiple reading measures collected in this study. The maze test is correlated significantly with reading comprehension (r=.70, p<.001), vocabulary (r= .64, p<.001), silent reading rate (r=.47, p<.001), ART (r= .56, p<.001), TOWRE SWE (r= .61, p<.001), and TOWRE PDE (r= .56, p<.001) on data from all readers. On data from poor readers, maze was significantly correlated with reading comprehension *(r*= .52, *p*<. 005), TOWRE SWE (*r*= .50, *p*<. 005) and TOWRE PDE (*r*= .39, *p*<.05). (The maze correlated -.04 with poor readers' RSPAN scores and .25 with all students' RSPAN scores, both unreliable.) These findings and previous research (Williams et al., in print) lend support to the use of maze as an effective measure of training effects in the cognitive processes of reading at the local level.

References

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods, 40*(1), 278-289.

Ackerman, B. P. (1984). Storage and processing constraints on integrating story information in children and adults. *Journal of experimental child psychology, 38*(1), 64-92.

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print* Cambridge: The MIT Press.

Albrecht, J. E., & Myers, J. L. (1995). Role of context in accessing distant information during reading. *Journal of experimental psychology. Learning, memory, and cognition, 21*(6), 1459.

Albrecht, J. E., & Myers, J. L. (1998). Accessing Distant Text Information during Reading: Effects of Contextual Cues. *Discourse processes, 26*(2-3), 87.

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: maintaining both local and global coherence. *Journal of experimental psychology. Learning, memory, and cognition, 19*(5), 1061.

Allinder, R. M., Dunse, L., Brunken, C. D., & Obermiller-Krolikowski, H. J. (2001). Improving fluency in at-risk readers and students with learning disabilities. *Remedial & Special Education, 22*(1), 48-54.

Allinder, R. M., Fuchs, L. S., & Fuchs, D. (1998). Curriculum-based measurement. In H. B. Vance (Ed.), *Psychological assessment of children: Best practices for school and clinical settings* (2nd ed., pp. 106-132). New York: Wiley.

Allington, R. L. (1983). Fluency: The neglected goal. *Reading Teacher, 36*(6), 556-561.

Allington, R. L. (1984). Content coverage and contextual reading in reading groups. *Journal of reading behavior, 16,* 85-96.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*(4), 369-406.

Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in reading and how children spend their time outside of school. *Reading research quarterly, 23,* 285-303.

Ardoin, S. P., McCall, M., & Klubnik, C. (2007). Promoting generalization of oral reading fluency: Providing drill versus practice opportunities. *Journal of behavioral education, 16,* 55-70.

August, D. L., Flavell, J. H., & Clift, R. (1984). Comparison of comprehension monitoring of skilled and less skilled readers. *Reading Research Quarterly, 20*(1), 39-53.

Baddeley, A., & Hitch, G. (1974). Working memory In G. H. Bower (Ed.), *The Psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York Academic Press

Baddeley, A., Logie, R., Nimmo-Smith, I., & Brereton, N. (1985). Components of fluent reading. *Journal of Memory and Language, 24*(1), 119-131.

Baker, L. (1979). Comprehension Monitoring: Identifying and Coping with Text Confusions. *Journal of reading behavior, 11*(4), 366-374.

Baker, L. (1985). Differences in the standards used by college students to evaluate their comprehension of expository prose. *Reading Research Quarterly, 20*(3), 297.

Baker, L. (1989). Metacognition, comprehension monitoring, and the adult reader. *Educational Psychology Review, 1*(1), 3-38.

Baker, L., & Anderson, R. I. (1982). Effects of inconsistent information on text processing: Evidence for comprehension monitoring. *Reading Research Quarterly, 17*(2), 281.

Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In R. B. P. D. Pearson, M. L. Kamil, & P. Mosenthal (Ed.), *Handbook of reading research* (pp. 353-394). New York: Longman.

Bereiter, C., & Bird, M. (1985). Use of thinking aloud in identification and teaching of reading comprehension strategies. *Cognition and instruction, 2*(2), 131.

Biemiller, A. (1977-1978). Relationships between oral reading rates for letters, words, and simple text in the development of reading achievement. *Reading Research Quarterly, 13*, 223-253.

Board of Regents. (2006). Regents' testing program passing rates for examinees repeating the test Fall semester, 2006 [Electronic Version]. Retrieved January 9, 2008, from httpp://www2.gsu.edu/~www.rtp/passing.htm

Breznitz, Z., & Share, D. (1992). Effects of accelerated reading rate on memory for text. *Journal of Educational Psychology, 84*(2), 193-199.

Brown, J. I., Fishco, V. V., & Hanna, G. (1993). *The Nelson-Denny Reading Test*. Chicago: Riverside.

Brown, J. I., Fishco, V. V., & Hanna, G. S. (1993). *Nelson Denny Reading Test: Forms G and H*. Itasca, IL: Riverside.

Cain, K., Oakhill, J. V., Hulme, C., & Joshi, R. M. (1998). Comprehension skill and inference-making ability: Issues of causality. In *Reading and spelling: Development and disorders.* (pp. 329-342). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Caplan, D., Waters, G., & Dede, G. (2007). Specialized working memory for language comprehension In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake & J. N. Towse (Eds.), *Variation in working memory* (pp. 272-302). Oxford Oxford University Press

Carbo, M. (1978). Teaching reading with talking books. *The Reading Teacher, 32*, 267-273.

Carver, R. P. (1990). *Reading rate : A review of research and theory*. San Diego: Academic Press.

Cattell, J. M. (1886). The time it takes to see and name objects *Mind, 11*, 63-65.

Chall, J. S. (1983). *Stages of reading development*. New York: McGraw-Hill.

Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The Reading Crisis: Why Poor Children Fall Behind?* . Cambridge: Harvard University Press.

Chomsky, C. (1978). When you still can't read in third grade. After decoding, what? In J. Samuels (Ed.), *What research has to say about reading instruction* (pp. 13-30). Neward, DE: International Reading Association

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*(3), 332-361.

Conte, R., & Humphreys, R. (1989). Repeated Readings Using Audiotaped Material Enhances Oral Reading in Children with Reading Difficulties. *Journal of communication disorders, 22*(1), 65-79.

Conway, A. R., & Engle, R. W. (1996). Individual differences in working memory capacity: more evidence for a general capacity theory. *Memory, 4*(6), 577.

Cook, A. E., Halleran, J. G., & O'Brien, E. J. (1998). What is readily available? A memory-based view of text processing. *Discourse processes, 26*, 109-129.

Creswell, J. W. (2003). *Research design : qualitative, quantitative, and mixed method approaches* (2nd ed.). Thousand Oaks, Calif.: Sage Publications.

Cunningham, A. E., Perry, E. K., & Stanovich, K. E. (2001). Converging evidence for the concept of orthographic processing. *Reading and writing, 14*(5/6), 549.

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of educational psychology, 83*, 264-274.

Cunningham, A. E., & Stanovich, K. E. (1993). Children's literacy environments and early word recognition subskills. *Reading and Writing: An interdisciplinary Journal, 5*, 193-204.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental psychology, 33*(6), 934-945.

Cunningham, A. E., & Stanovich, K. E. (1998). The impact of print exposure on word recognition. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 235-262). Mahwah: Lawrence Erlbaum Associates.

Curtis, M. (1980). Development of components of reading skill. *Journal of Educational Psychology, 72*(5), 656-669.

Dahl, P. R. (1974). *An Experimental Program for Teaching High Speed Word Recognition and Comprehension Skills*.

Dahl, P. R. (1979). An experimental program for teaching high speed word recognition and comprehension skills. In J. E. Button, T. C. Lovitt & T. D. Rowland (Eds.), *Communications research in learning disabilities and mental retardation* Baltimore: University Park Press

Daly, E. J., Chafouleas, S. M., & Skinner, C. H. (2004). *Interventions for reading problems: Designing and evaluating effective strategies. Practical intervention in the schools series*. New York: Guilford Press.

Daneman, M. (1987). Reading and working memory. In J. R. Beech & A. M. Colley (Eds.), *Cognitive approaches to reading* (pp. 57-80). Chichester: John Wiley & Sons.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior, 19*(4), 450.

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of experimental psychology. Learning, memory, and cognition, 9*(4), 561.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422-433.

Daneman, M., & Tardif, T. (1987). Working memory and reading skill re-examined. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 491-508). London: Erlbaum.

Dell, G. S., McKoon, G., & Ratcliff, R. (1983). The activation of antecedent information during the processing of anaphoric reference in reading. *Journal of Verbal Learning & Verbal Behavior, 22*(1), 121-132.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*(219-232).

Deno, S. L., & Marston, D. (2006). Curriculum-based measurement of oral reading: An indicator of growth in fluency. In J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 179-203). Neward, DE International Reading Association

Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading *Exceptional Children, 49*, 36-45.

Dowhower, S. L. (1987). Effects of Repeated Reading on Second-Grade Transitional Readers' Fluency and Comprehension. *Reading Research Quarterly, 22*(4), 389.

Durkin, D. (1978). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly, 14*(4), 481.

Durkin, D. (1993). *Teaching them to read* (6th ed.). Boston: Allyn and Bacon.

Dwyer, E. J., & West, R. F. (1989). Demystifying "Speed Reading": A Practical Approach for Increasing Rate. *Forum for Reading, 21*(1), 68-74.

Dwyer, E. J., & West, R. F. (1994). *Effects of Sustained Silent Reading on Reading Rate among College Students* o. Document Number)

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of research in reading, 18*(2), 116-125.

Ehri, L. C. (2005). Learning to Read Words: Theory, Findings, and Issues. *Scientific Studies of Reading, 9*(2), 167.

Ehri, L. C., & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading & writing quarterly: Overcoming learning difficultuies, 14*, 135-164.

Ehri, L. C., & Wilce, L. S. (1983). Development of word identification speed in skilled and less skilled beginning readers. *Journal of educational psychology, 75*(1), 3-18.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: a test of four hypotheses. *Journal of experimental psychology. Learning, memory, and cognition, 18*(5), 972.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological review, 102*(2), 211.

Fleisher, L. S., Jenkins, J. R., & Pany, D. (1979). Effects on Poor Readers' Comprehension of Training in Rapid Decoding. *Reading Research Quarterly, 15*(1), 30.

Fletcher, C. R. (1994). Levels of representation in memory for discourse. In M. A. Gernsbacher (Ed.), *Handbook of pyscholinguistics* (pp. 589-607). San Diego: Academic Press.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116-124.

Friedman, N. P., & Miyake, A. (2005). Comparison of four scoring methods for the reading span test. *Behavior Research Methods, 37*(4), 581.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, J. Marshall & M. Coltheart (Eds.), *Surface dyslexia* (pp. 301-330). London: Erlbaum.

Fuchs, L. S., & Fuchs, D. (1992). Identifying a measure for monitoring student reading progress. *School Psychology Review, 21*(1), 45-58.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement, using a reading maze task. *Exceptional Children, 58*(5), 436-450.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*(3), 239-256.

Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education, 9*(2), 20-29.

Garner, R. (1981). Monitoring and Resolving Comprehension Obstacles: An Investigation of Spontaneous Text Lookbacks among Upper-Grade Good and Poor Comprehenders. *Reading Research Quarterly, 16*(4), 569.

Garner, R., & Reis, R. (1981). Monitoring and resolving comprehension obstacles: An investigation of spontaneous text lookbacks among upper-grade good and poor comprehenders. *Reading Research Quarterly, 16*(4), 569-582.

Garnham, A., Oakhill, J. V., & Johnson-Laird, P. N. (1982). Referential continuity and the coherence of discourse. *Cognition, 11*, 29-46.

Gernsbacher, M. A. (1990). *Language comprehension as structure building* Hillsdale, NJ: Lawrence Erlbaum Associates.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of experimental psychology. Learning, memory, and cognition, 16*(3), 430-445.

Gerrig, R. J., & O'Brien, E. J. (2005). The scope of memory-based processing *Discourse processes, 39*(2&3), 225-242.

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of memory and language, 31*(2), 129-151.

Gough, P. B., Hoover, W. A., & Peterson, C. L. (1996). Some observations on a simple view of reading. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 1-14). Mahwah, NJ: Lawrence Erlbaum Associates.

Graesser, A. C., Singer, M., & Trabasso, T. T. (1994). Constructing inferences during narrative text comprehension. *Psychological review, 101*(3), 371.

Guthrie, J. T. (1973). Models of reading and reading disability. *Journal of Educational Psychology, 65*, 9-18.

Guthrie, J. T., Seifert, M., Burnham, N. A., & Caplan, R. I. (1974). The maze technique to assess, monitor reading comprehension. *Reading Teacher, 28*(2), 161-168.

Guthrie, J. T., Wigfield, A., Metsala, J. L., & Cox, K. E. (1999). Motivational and cognitive predictors of text comprehension and reading amount. *Scientific studies of reading, 3*(3), 231-256.

Hakala, C. M., & O'Brien, E. J. (1995). strategies for resolving coherence breaks in reading *Discourse Processes, 20,* 167-185.

Hannon, B., & Daneman, M. (2001). A New Tool for Measuring and Understanding Individual Differences in the Component Process of Reading Comprehension. *Journal of educational psychology, 93*(1), 103-128.

Hannon, B., & Daneman, M. (2004). Shallow semantic processing of text: An individual-differences account. *Discourse processes, 37*(3), 187 - 204.

Harris, A. J., & Sipay, E. R. (1990). *How to increase reading ability: A guide to developmental & remedial methods* (9th ed.). New York: Longman.

Harris, P. L., Kmithof, A., Terwogt, M. M., & Visser, T. (1981). Children's Detection and Awareness of Textual Anomaly. *Journal of experimental child psychology, 31*(2), 212.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children* Baltimore: Brookes Publishing Company.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*(3), 356-388.

Hayes, D. P., & Ahrens, M. (1988). Vocabulary simplification for children: A special case of 'motherese'? *Journal of Child Language 15* 395-410.

Herman, P. A. (1985). The Effect of Repeated Readings on Reading Rate, Speech Pauses, and Word Recognition Accuracy. *Reading research quarterly, 20*(5), 553.

Hiebert, E. H. (2005b). The effects of text difficulty on second graders' fluency developtment. *Reading psycholgoy, 26*, 1-27.

Hiebert, E. H., & Fisher, C. W. (2005). A review of the national reading panel's studies on fluency: On the role of text. *The Elementary School Journal, 105*, 443-460.

Hirsch, E. D. (2003). Reading comprehension requires knowledge of words and of the world. *American Educator, 27*, 10-29, 44-45.

Hollingsworth, P. M. (1978). An experimental approach to the impress method of teaching reading *The Reading Teacher, 31*, 624-626.

Homan, S. P., Klesius, J. P., & Hite, C. (1993). Effects of repeated readings and nonrepetitive strategies on students' fluency and comprehension. *Journal of Educational research, 87*(2), 94-99.

Hoover, W. A., & Gough, P. B. (1990a). The simple view of reading. *Reading and writing, 2*(2), 127.

Hoover, W. A., & Gough, P. B. (1990b). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2*, 127-160.

Huey, E. B. (1908/1968). *The psychology and pedagogy of reading* Boston: MIT Press.

Jackson, M. D., & McClelland, J. L. (1979). Processing determinants of reading speed. *Journal of Experimental Psychology: General, 108*(2), 151-181.

Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology, 95*(4), 719-729.

Jenkins, J. R., & Jewell, M. (1993). Examining the validity of two measures for formative teaching: Reading aloud and maze. *Exceptional Children, 59*(5), 421-432.

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of educational psychology, 78*, 243-255.

Juel, C., & Holmes, B. (1981). Oral and silent reading of sentences. *Reading Research Quarterly, 16*(4), 545-568

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological review, 99*(1), 122.

Kennedy, C. H. (2005). *Single-case designs for educational research*. Boston: Pearson/A & B.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*, 580-602.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, N.J.,: Lawrence Erlbaum Associates; distributed by Halsted Press Division, Wiley, New York.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological review, 95*(2), 163.

Kintsch, W. (2004). The construction-integration model of text comprehension and its implications for instruction. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (pp. 1270-1328). Newark, DE: International Reading Association.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological review, 85*(5), 363.

Kowal, S., O'Connell, D. C., O'Brien, E. A., & Bryant, E. T. (1975). Temporal aspects of reading aloud and speaking: Three experiments. *American Journal of Psychology, 88*(4), 549-569.

Krashen, S. D. (1993). *The power of reading : insights from the research*. Englewood, Colo.: Libraries Unlimited.

Kuhn, M. R. (2005a). A comparative study of small group fluency instruction. *Reading Psychology, 26*(2), 127-146.

Kuhn, M. R. (2005b). A Comparative Study of Small Group Fluency Instruction. *Reading Psychology, 26*(2), 127.

Kuhn, M. R., Schwanenflugel, P. J., Morris, R. D., Morrow, L. M., Woo, D. G., Meisinger, E. B., et al. (2006). Teaching children to become fluent and automatic readers. *Journal of Literacy Research, 38*(4), 357-387.

Kuhn, M. R., & Stahl, S. A. (2000). Fluency: A Review of Developmental and Remedial Practices. 47.

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95*(1), 3.

Kuhn, M. R., Stahl, S. A., & Center for the Improvement of Early Reading Achievement, A. A. M. I. (2000). *Fluency: A Review of Developmental and Remedial Practices*.

Laberge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading *Cognitive Psychology, 6*, 293-323.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*(2), 211-240.

Lee, Y., Lu, M., & Ko, H. (2007). Effects of skill training on working memory capacity. *Learning and Instruction, 17,* 336-344.

Liberman, I. Y., Shankweiler, D. P., Fischer, F., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of experimental child psychology, 18*, 201-212.

Logan, G. D. (1978). Attention in character-classification tasks: Evidence for the automaticity of component stages. *Journal of Experimental Psychology: General, 107*(1), 32-63.

Logan, G. D. (1979). On the use of a concurrent memory load to measure attention and automaticity. *Journal of Experimental Psychology: Human Perception and Performance, 5*(2), 189-207.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*(4), 492.

Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 13*(2), 123-146.

Long, D. L., & Chong, J. L. (2001). Comprehension skill and global coherence: A paradoxical picture of poor comprehenders' abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(6), 1424-1429.

Long, D. L., Johns, C. L., & Morris, P. E. (2006). Comprehension ability in mature readers. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of Psycholinguistics* (2 ed., pp. 801-833). Amsterdam: Elsevier.

Long, D. L., Oppy, B. J., & Seely, M. R. (1994). Individual differences in the time course of inferential processing. *Journal of experimental psychology. Learning, memory, and cognition, 20*(6), 1456.

Long, D. L., Oppy, B. J., & Seely, M. R. (1997). Individual differences in readers' sentence- and text-level representations. *Journal of memory and language, 36,* 129-145.

Long, D. L., Seely, M. R., & Oppy, B. J. (1999). The strategic nature of less skilled readers' suppression problems. *Discourse processes, 27*(3), 281.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: A comment on Just & Carpenter (1992) and Waters & Caplan (1996). *Psychological Review, 109*, 35-54.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *cognitive psychology, 24*, 56-98.

MacKay, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review, 89*(5), 483-506.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological bulletin, 109*(2), 163-203.

Markman, E. M. (1977). Realizing that you don't understand: A preliminary investigation. *Child development, 48*(3), 986-992.

Markman, E. M. (1979). Realizing that you don't understand: elementary school children's awareness of inconsistencies. *Child development, 50*(3), 643-655.

Markman, E. M. (1981). Comprehension monitoring. In W. P. Dickson (Ed.), *Children's oral communication skills* (pp. 61-84). New York: Academic Press.

Marston, D. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. In M. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18-78). New York: Guilford.

Mathes, P. G. (1993). Peer-Mediated Reading Instruction in Special Education Resource Rooms. *Learning Disabilities Research & Practice, 8*(4), 233.

Mathes, P. G., & Fuchs, L. S. (1993). Peer-Mediated Reading Instruction in Special Education Resource Rooms. *Learning Disabilities Research & Practice, 8*(4), 233.

McKoon, G., Gerrig, R. J., & Greene, S. B. (1996). Pronoun resolution without pronouns: some consequences of memory-based text processing. *Journal of experimental psychology. Learning, memory, and cognition*, *22*(4), 919-932.

McKoon, G., & Ratcliff, R. (1992). Inference During Reading. *Psychological Review, 99*(3), 440-466.

Mercer, C. D., Campbell, K. U., Miller, M. D., Mercer, K. D., & Lane, H. B. (2000). Effects of a reading fluency intervention for middle schoolers with specific learning disabilities. *Learning Disabilities Research & Practice, 15*, 179-189.

Mertens, D. M. (1998). *Research methods in education and psychology : integrating diversity with quantitative & qualitative approaches*. Thousand Oaks, Calif.: Sage Publications.

Meyer, M. S., & Felton, R. H. (1999). Repeated reading to enhance fluency: Old approaches and new directions. *Annals of dyslexia, 49*, 283-306.

Miller, J., & Schwanenflugel, P. J. (2006). Prosody of Syntactically Complex Sentences in the Oral Reading of Young Children. *Journal of educational psychology, 98*(4), 839-853.

Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of educational psychology, 94*(2), 249.

Morrow, D. G., Greenspan, S. L., & Bower, G. H. (1987). Accessibility and situation models in narrative comprehension. *Journal of memory and language, 26*(2), 165-187.

Moyer, S. S. B. (1982). Repeated reading. *Journal of learning disabilities, 15*(10), 619-623.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the Discourse Representation during Reading. *Discourse processes, 26*(2-3), 131.

Myers, J. L., O'Brien, E. J., Albrecht, J. E., & Mason, R. A. (1994). Maintaining global coherence during reading. *Journal of experimental psychology. Learning, memory, and cognition, 20*(4), 876.

Nagy, W., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly, 19*, 304-330.

Nagy, W., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal, 24*(2), 237.

Nathan, R. G., & Stanovich, K. E. (1991). The causes and consequences of differences in reading fluency. *Theory into practice, XXX*(3), 176-183.

Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of experimental child psychology, 73*(2), 139-158.

National Assessment of Educational Progress. (2003). *2002, 1998, 1994 and 1992 Reading assessments*: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

National Center for Educational Statistics. (2001). *The national assessment of educational progress (NAEP)*. Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Institute of Child Health and Human Development. (2000a). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, D.C.: National Institute of Child Health and Human Development, National Institutes of Health.

National Institute of Child Health and Human Development. (2000b). *Teaching children to read: An evidence-based assessment of the scientific literature on reading and its implications for reading instruction*. Washington, DC: U.S. Government Printing Office.

Neuman, S. B., & Celano, D. (2001). Access to print in low-income and middle-income communities: An ecological study of four neighborhoods. *Reading Research Quarterly, 36*(1), 8-26.

Newsom, S. (1979). Rock'n roll'n reading. *Journal of Reading, 22*, 726-730.

No Child Left Behind Act. (2002). No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425.

Norman, D. A. (1968). Toward a theory of memory and attention. *Psychological review, 75,* 522-536.

O'Brien, E. J., & Albrecht, J. E. (1992). Comprehension strategies in the development of a mental model. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 18*(4), 777.

O'Brien, E. J., Plewes, P. S., & Albrecht, J. E. (1990). Antecedent retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(2), 241-249.

O'Brien, E. J., Rizzella, M. L., Albrecht, J. E., & Halleran, J. G. (1998). Updating a situation model: A resonance text processing view. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1200-1210.

O'Connor, R. E., White, A., & Swanson, L. H. (2007). Repeated reading versus continuous reading: Influences on reading fluency and comprehension. *Exceptional Children, 74*(1), 31-46.

O'Shea, L. J., Sindelar, P. T., & O'Shea, D. J. (1985). The effects of repeated readings and attentional cues on reading fluency and comprehension. *Journal of Reading Behavior, 17*(2), 129-142.

Oakhill, J. V. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology, 54,* 31-39.

Oakhill, J. V., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and writing*, *18*(7-9), 657.

Oakhill, J. V., & Yuill, N. M. (1986). Pronoun resolution in skilled and less-skilled comprehenders: Effects of memory load and inferential complexity. *Language & Speech*, *29*(1), 25-37.

Oakhill, J. V., Yuill, N. M., & Parkin, A. (1986). On the nature of the difference between skilled and less-skilled comprehenders. *Journal of Research in Reading, 9*(2), 80.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience, 7*, 75-79.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring Activities. *Cognition and Instruction, 1*(2), 117.

Palmer, J., Mcleod, C. M., Hunt, E., & Davidson, J. E. (1985). Information processing correlates of reading. *Journal of Memory and Language, 24*(1), 59-88.

Park, H. M. (2008). Univariate Analysis and Normality Test Using SAS, Stata, and SPSS. The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing, Indiana University.

Pearlmutter, N. J., & MacDonald, M. C. (1995). Individual differences and probabilistic constraints in syntactic ambiguity resolution *Journal of memory and language, 34*, 521-542.

Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383.

Perfetti, C. A., & Britt, M. A. (1995). Where do propositions come from? . In C. A. Weaver, S. Mannes & C. R. Fletcher (Eds.), *Discourse comprehension: Essays in honor of Walter Kintsch* (pp. 11-34). Hillsdale, NJ: Lawrence Erlbaum Associates.

Perfetti, C. A., & Goldman, S. R. (1974). Thematization and sentence retrieval. *Journal of Verbal Learning and Verbal Behavior, 13*, 70-79.

Perfetti, C. A., & Goldman, S. R. (1976). Discourse memory and reading comprehension skill. *Journal of Verbal Learning and Verbal Behavior, 14*, 33-42.

Perfetti, C. A., & Hogaboam, T. W. (1975). Relationship Between Single Word Decoding and Reading Comprehension Skill.

Perfetti, C. A., & Lesgold, A. (1977). Discourse comprehension and sources of individual differences In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* Hillsdale, NJ: Erlbaum

Perfetti, C. A., Marron, M. A., & Foltz, G. (1996). Sources of comprehension failures: Theoretical perspectives and case studies In C. Cornoldi & J. Oakhill (Eds.), *Reading Comprehension Difficulties: Processes and Intervention* (pp. 137-165). Mahwah, NJ: Lawrence Erlbaum Associates.

Pikulski, J. J. (2006). Fluency: A developmental and language perspective. In J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 70-93). Newark, DE: International Reading Association.

Pikulski, J. J., & Chard, D. J. (2005). Fluency: Bridge Between Decoding and Reading Comprehension. *The Reading teacher, 58*(6), 510.

Pinnell, G. S., Pikulski, J. J., Wixson, K. K., Campbell, J. R., Gough, P. B., & Beatty, A. S. (1995). *Listening to children read aloud*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Posner, M. I., Lewis, J. L., & Conrad, C. (1972). Component processes in reading:Aperformance analysis. In J. F. Kavanaugh & I. G. Mattingly (Eds.), *Language by ear and by eye* (pp. 159-192). Cambridge, MA: MIT Press.

Potts, G. R., Keenan, J. M., & Golding, J. M. (1988). Asessing the occurrence of elaborative inferences: Lexical decision versus naming *Journal of memory and language, 27*, 399-415.

Potts, G. R., & Peterson, S. B. (1985). Incorporation versus compartmentalization in memory for discourse. *Journal of memory and language, 24*(1), 107-118.

Pressley, M., Hilden, K. R., & Shankland, R. K. (2006). *An evaluation of end-grade-3 Dynamic Indicators of Basic Early Literacy Skills (DIBELS): Speed reading without comprehension, predicting little*. East Lansing: Michigan State University, College of Education, Literacy Achievement Research Centero. Document Number)

Rashotte, C. A., & Torgesen, J. K. (1985). Repeated reading and reading fluency in learning disabled children. *Reading Research Quarterly, 20*(2), 180.

Rasinski, T. V. (1989). *The Effects of Repeated Reading and Repeated Listening while Reading on Reading Fluency* o. Document Number)

Rasinski, T. V., Padak, N. D., McKeon, C. A., Wilfong, L. G., Friedauer, J. A., & Heim, P. (2005). Is reading fluency a key for successful high school reading? *Journal of Adolescent & Adult Literacy, 49*(1), 22-27.

Ratcliff, R., & McKoon, G. (1978). Priming in item recognition: Evidence for the propositional structure of sentences. *Journal of verbal learning and verbal behavior, 17*(4), 403.

Reitsma, P. (1983). Printed Word Learning in Beginning Readers. *Journal of experimental child psychology, 36*(2), 321.

Research & Education Association. (2007). *Reading comprehension builder for standardized tests* Research & Education Association

Reutzel, D. R., & Hollingsworth, P. M. (1991). Reading time in school: Effect on fourth graders' performance on a criterion-referenced comprehension test. *Journal of Educational research, 84*, 171-176.

Rizzella, M. L., & O'Brien, E. J. (1996). Accessing global causes during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(5), 1208-1218.

Robinson, E. J., & Robinson, W. P. (1977a). Children's explanation of communication failure and the inadequacy of the misunderstood message. *Developmental Psychology, 13*, 156-161.

Romano, L. (2005, December 25, 2005). Literacy of college graduates is on decline: Survey's finding of a drop in reading proficiency is inexplicable, experts say. *The Washington Post*, p. A12,

Rosenbloom, P. S. (1984). *The chunking of goal hierarchies: A model of practice and stimulus-response compatibility.* ProQuest Information & Learning, US.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of educational research, 64*(4), 479.

Ruffman, T. (1996). Reassessing children's comprehension-monitoring skills. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and Intervention* (pp. 33-68). Mahwah, NJ: Lawrence Erlbaum Associates.

Samuels, S. J. (1979). The method of repeated reading. *The Reading Teacher, 12*, 177-186.

Samuels, S. J. (2007). The DIBELS tests: Is speed of barking at print what we mean by reading fluency? *Reading Research Quarterly, 42*(4), 563-566.

Schiller, P. P. H. (1966). Developmental study of color-word interference. *Journal of Experimental Psychology, 72*(1), 105-108.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*(1), 1-66.

Schreiber, P. A. (1980). On the acquisition of reading fluency. *Journal of Reading Behavior, 12*, 177-186.

Schreiber, P. A. (1991). Understanding prosody's role in reading acquisition. *Theory into practice, XXX*(3), 158-164.

Schwanenflugel, P. J., Hamilton, A. M., Kuhn, M. R., Wisenbaker, J. M., & Stahl, S. A. (2004). Becoming a fluent reader: Reading skill and prosodic features in the oral reading of young readers. *Journal of educational psychology, 96*(1), 119-129.

Schwanenflugel, P. J., Meisinger, E. B., Wisenbaker, J. M., Kuhn, M. R., Strauss, G. P., & Morris, R. D. (2006). Becoming a Fluent and Automatic Reader in the Early Elementary School Years. *Reading Research Quarterly, 41*(4), 496.

Seigneuric, A., Ehrlick, M., Oakhill, J. V., & Yuill, N. M. (2000). Working memory resources in children's reading comprehension. *Reading & Writing: An Interdisciplinary Journal 13*, 81-103.

Shah, P. P., & Miyake, A. A. (1996). The separability of working memory resources for spatial thinking and language processing: an individual differences approach. *Journal of experimental psychology. General, 125*(1), 4.

Shankweiler, D. P., & Crain, S. S. (1986). Language mechanisms and reading disorder: a modular approach. *Cognition, 24*(1-2), 139.

Shankweiler, D. P., Lundquist, E., Katz, L., Stuebing, K. K., Fletcher, J. M., & Brady, S. (1999). Comprehension and decoding: Patterns of association in children with reading difficulties. *Scientific Studies of Reading, 3*(1), 69-94.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*, 151-218.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*(2), 127-190.

Singer, M. (1993). Causal bridging inferences: Validating consistent and inconsistent sequences. *Canadian Journal of Experimental Psychology, 47*, 340-359.

Singer, M., Andrusiak, P., Reisdorf, P., & Black, N. L. (1992). Individual differences in bridging inference processes. *Memory & cognition, 20*(5), 539.

Singer, M., & Ferreira, F. (1983). Inferring consequences in story comprehension. *Journal of Verbal Learning & Verbal Behavior, 22*(4), 437-448.

Singer, M., & Halldorson, M. (1996). Constructing and Validating Motive Bridging Inferences. *Cognitive Psychology, 30*(1), 1.

Singer, M., & Ritchot, K. F. M. (1996). The role of working memory capacity and knowledge access in text inference processing. *Memory & cognition, 24*(6), 733-743.

Slocum, T. A., Street, E. M., & Gilberts, G. (1995). A review of research and theory on the relation between oral reading rate and reading comprehension. *Journal of Behavioral Education, 5*(4), 377.

Snow, C. E., Barnes, W. S., Chandler, J., Goodman, I. F., & Hemphill, L. (1991). *Unfulfilled Expectations: Home and School Influences on Literacy*. Cambridge, MA: Harvard University Press.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Spargo, E., & Williston, G. R. (1975). *Timed Readings: Level 8*. Providence, R. I.: Jamestown Publishers.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-407.

Stanovich, K. E., & Cunningham, A. E. (1993). Where does knowledge come from? Specific associations between print exposure and information acquisition. *Journal of educational psychology, 85*(2), 211-229.

Stanovich, K. E., Cunningham, A. E., & Feeman, D. J. (1984). Intelligence, cognitive skills and early reading progress. *Reading research quarterly, 19*, 278-303.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24*(4), 402-433.

Stecker, S. K., Roser, N. L., & Martinez, M. G. (1998). Understanding oral reading fluency. In T. Shanahan & F. V. Rodriguez-Brown (Eds.), *47th yearbook of the National Reading Conference* (pp. 295-310). Chicago: National Reading Conference.

Sternberg, R. J., & Powell, J. S. (1983). Comprehending Verbal Comprehension. *The American psychologist, 38*(8), 878.

Stevens, J. P. (1999). *Intermediate statistics: A modern approach* (2 ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Stockburger, D. (2001). *Introductory Statistics: Concepts, Models and Applications, 2e.* .

Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children: the role of language comprehension and working memory skills. *Reading and writing, 4*(3), 245-256.

Strong, A. C., Wehby, J. H., Falk, K. B., & Lane, K. L. (2004). The impact of a structured reading curriculum and repeated reading on the performance of junior high students with emotional and behavioral disorders. *School Psychology Review, 33*(4), 561-581.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)Consideration of context effects. *Journal of verbal learning and verbal behavior, 18*, 523-534.

Tabachnick, B. G., & Fidell, L. S. (2005). *Using multivariate statistics* (5th ed.). Boston: Allyn and Bacon.

Therrien, W. J. (2004a). Fluency and Comprehension Gains as a Result of Repeated Reading. *Remedial and special education, 25*(4), 252.

Therrien, W. J. (2004b). Fluency and comprehension gains as a result of repeated reading: A meta-analysis. *Remedial and special education, 25*(4), 252.

Therrien, W. J., & Kubina, R. M. (2006). Developing Reading Fluency with Repeated Reading. *Intervention in school and clinic, 41*(3), 156.

Till, R. E., Mross, E. F., & Kintsch, W. (1988). Time course of priming for associate and inference words in a discourse context. *Memory & Cognition, 16*(4), 283-298.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of Word Reading Efficiency*. Austin, TX: PRO-ED Publishing, Inc.

Tzelgov, J., Henik, A., & Leiser, D. (1990). Controlling Stroop interference: Evidence from a bilingual task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(5), 760-771.

van den Broek, P., Risden, K., & Husebye-Hartmann, E. (1995). The role of readers' standards for coherence in the generation of inferneces during reading. In R. F. Lorch & E. J. O'brien (Eds.), *Sources of coherence in reading* (pp. 353-374). Hillsdale, NJ: Lawrence Erlbaum Associates.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Vosniadou, S., Pearson, P. D., & Rogers, T. (1988). What causes children's failures to detect inconsistencies in test? Representation versus comparison difficulties. *Journal of educational psychology, 80*(1), 27.

Wagoner, S. A. (1983). Comprehension Monitoring: What It Is and What We Know about It. *Reading Research Quarterly, 18*(3), 328.

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *The Quarterly journal of experimental psychology. A, Human experimental psychology, 49*(1), 51-75.

Waters, G. S., & Caplan, D. (1996b). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Pyschological Review 103*, 761-772.

Waters, G. S., & Caplan, D. (2004). Verbal working memory and online syntactic processing: Evidence from self-paced listening *Quarterly Journal of Experimental Psychology, 57*, 129-164.

Waters, G. S., Caplan, D., & Hildebrandt, N. (1987). Working memory and written sentence comprehension In M. Coltheart (Ed.), *Attention and Performance XII: The Psychology of Reading* (pp. 531-555). Hillsdale, NJ: Erlbaum

Williams, R. S., Ari, O., & Santamaria, C. N. (in print). Measuring college students' reading comprehension ability using cloze tests. *Journal of Research in Reading*.

Winograd, P. N., & Johnston, P. (1982). Comprehension monitoring and the error detection paradigm. *Journal of reading behavior, 14*(1), 61-76.

Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific studies of reading, 5*(3), 211-239.

Woodcock, R. (1998). *The Woodcock Reading Mastery Tests-NU*. Circle Pines, MN: American Guidance Service.

Young, A. R., Bowers, P. G., & MacKinnon, G. E. (1996). Effects of prosodic modeling and repeated reading on poor readers' fluency and comprehension. *Applied psycholinguistics, 17*(1), 59-84.

Yuill, N. M., & Oakhill, J. V. (1991). *Children's problems in text comprehension: An experimental investigation*.

Yuill, N. M., Oakhill, J. V., & Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British Journal of Psychology*, *80* (3), 351-361.

Zabrucky, K., & Moore, D. (1989). Children's Ability to Use Three Standards to Evaluate Their Comprehension of Text. *Reading Research Quarterly*, *24*(3), 336.

APPENDIXES

APPENDIX A

Example Passage from Long and Chong (2001, p. 1429)

Introduction
Ken and his friend Mike had been looking for summer hobbies for quite some time. They were both college professors and they had the summers off from teaching. This meant that they both had plenty of time to try new things.
First-Character-Consistent Elaboration
Ken was a big man and always tried to keep in shape by jogging and lifting weights. His 250-pound body was solid muscle. Ken loved tough physical contact sports, which allowed him to match his strength against another person.
First-Character-Inconsistent Elaboration
Ken was a small man and didn't worry about staying in shape. His small 120-pound body was all skin and bones. Ken hated contact sports, but enjoyed noncontact sports, such as golf and bowling which he could practice along.
Second-Character-Consistent Elaboration
Mike was a big man and always tried to keep in shape by jogging and lifting weights. His 250-pound body was solid muscle. Mike loved tough physical contact sports, which allowed him to match his strength against another person.
Second-Character-Inconsistent Elaboration
Mike was a small man and didn't worry about staying in shape. His small 120-pound body was all skin and bones. Mike hated contact sports, but enjoyed noncontact sports, such as golf and bowling which he could practice along.
Filler—Global Coherence Condition
While walking downtown during their lunch break one day, Ken and Mike passed a new gymnasium. They noticed the display in the window. It was an advertisement for the gym's summer sports program. They started looking at the advertisement and were impressed with the long list of activities that the gym sponsored. As they continued to look over the list, they became very excited. It seemed interesting so Ken and Mike went inside.
Filler—Local Coherence Condition
While walking by a new gymnasium downtown, Ken and Mike saw a flyer for the gym's summer sports program.
Target Sentence
Ken decided to enroll in boxing classes.
Posttarget Sentence
He felt this would be the perfect hobby.
Comprehension Question (Experiment 1)
Was Ken looking for a hobby?
Verification Probe (Experiment 2)
[Ken/Mike] liked noncontact sports.

APPENDIX B

Study Brochure

Please fill out the following information and return it to the researcher

Name: ....

Phone: ....

Email: ....

Contact
Omer Ari
ariomer@hotmail.com
404 422 7446

PROJECT RIFLE

Reading **F**luency **I**nterventions **f**or **L**ow Comprehenders **E**nrolled in College

PROJECT RIFLE

Project RIFLE was designed on the basis of an observed pattern on required reading tests. Every semester hundreds of undergraduates take various reading tests (e.g., Regents' Reading Test, Compass, Regents' OSTEP, etc). Some of these students demonstrate significant difficulty passing the test. **Several studies have pointed out that their difficulty may be due to low fluency skills.**

What is fluency?

Fluency refers to reading correctly, quickly, and with expression. According to previous research, **fluency is strongly related to reading comprehension success.**

Why this study?

Unlike previous studies, **this study will examine students' ability to monitor their comprehension at the end of a fluency-training program.**

Why should you participate?

If you experience difficulty passing a reading test or can't finish your reading assignments in time, you should consider participating in this study.

Participation may improve:

- your reading speed;
- your reading comprehension;
- your ability to read unfamiliar words;
- your ability to monitor the state of your comprehension;
- your ability to fix a comprehension break when it occurs.

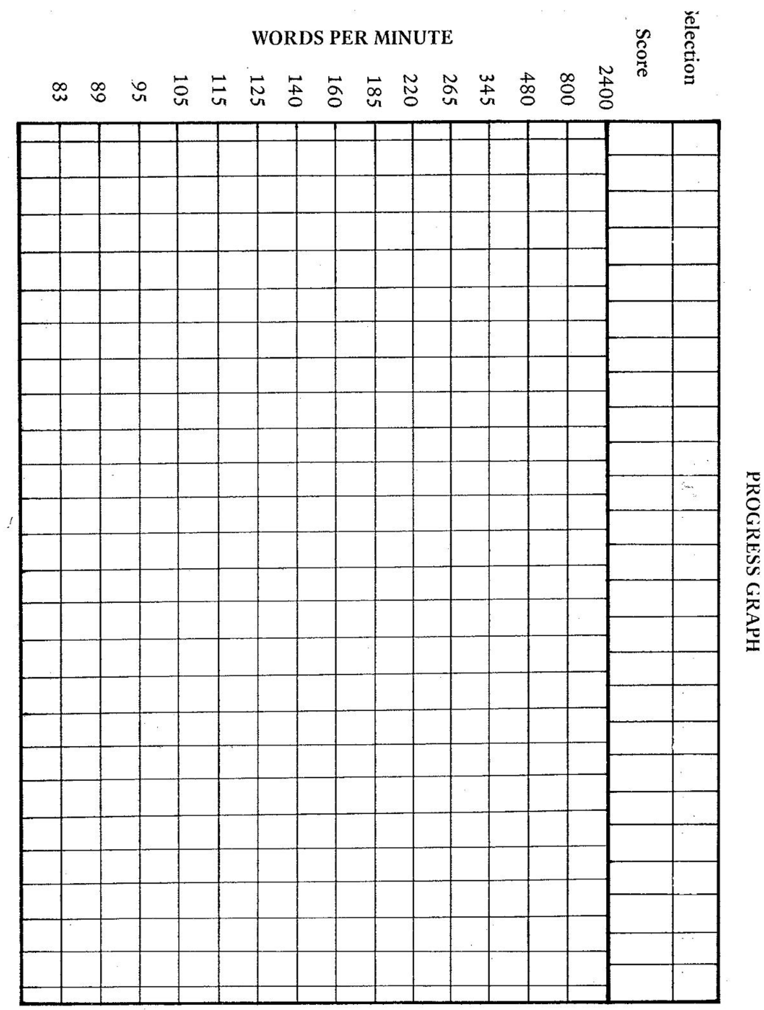
How long is the training?

The study will last seven weeks. Participants will complete fluency training in 30-minute sessions three times a week.

PROJECT RIFLE

APPENDIX C

Progress Graph



118

APPENDIX D

Time-Rate Conversions Table

73

TIME-RATE CONVERSIONS

Reading Time	Words per Minute	Reading Time	Words per Minute
:10	2400	2:40	150
:20	1200	2:50	140
:30	800	3:00	135
:40	600	3:10	125
:50	480	3:20	120
1:00	400	3:30	115
1:10	345	3:40	110
1:20	300	3:50	105
1:30	265	4:00	100
1:40	240	4:10	95
1:50	220	4:20	92
2:00	200	4:30	89
2:10	185	4:40	85
2:20	170	4:50	83
2:30	160	5:00	80

8a
PB 199
120

APPENDIX E

Sample *Timed Readings* Passage

Island on the Move (Grade Level 4)

Puerto Rico is neither a state, colony, nor territory. Since July 1952, it has been a free commonwealth or a free associated state of the United States. Puerto Ricans enjoy a representative form of government. The legislative branch is headed by the elected governor, assisted by a cabinet. It also has its own court system.

The flavor of Puerto Rico is Spanish, for the island was a Spanish colony from 1508 to 1898, when it was ceded to the United States as a result of the Spanish-American War.

A land reform program, which started about 25 years ago, has given small plots to thousands of squatter families. Yearly profits of plantations controlled by a government agency are distributed among all employees. Redevelopment projects have transformed San Juan. Waterfront shacks and slum neighborhoods have given way to neat housing projects and modern resort hotels complete with pools, beaches, and shops.

Along with economic advances, a cultural revival program has been carried on under government control. This program is aimed at keeping Puerto Rico's cultural heritage and developing contemporary arts and culture. As a result of this program, one can examine treasures in 16^{th} and 17^{th} century churches and museums, and attend fiestas and ceremonies in which the old customs and costumes have been revived.

Elementary education has been free and required since 1899. The literacy rate is high – about 88 percent for those 10 years and older. There are four schools of learning on the island.

English is taught in the schools, though the instruction is in Spanish below the senior high school level. In San Juan, is about as common as Spanish.

San Juan is Puerto Rico's modern capital and home of nearly 500,000 people. It is built around the busiest port in the West Indies. The port and the large, modern airport make San Juan the crossroads of North and South American traffic.

The oldest part of San Juan, now only six blocks square, was built and fortified nearly 500 years ago. Large bridges connect it to the main island. *Plaza Colon* is the center of the old city and the location of the bus terminal.

Balconies protected by wrought iron overhang the narrow streets of the old city. Strolling down these streets can give the impression of being transported into old Spain.

Comprehension Questions

Literal Questions (Recalling Facts)

1. Puerto Rico is considered
   a. a city.
   b. a territory.
   c. a commonwealth.

2. The executive branch of the Puerto Rican government is headed by
   a) President
   b) an ambassador
   c) a governor

3. Puerto Rico was once controlled by
   a. Portugal.
   b. Spain.
   c. France.

4. The literacy rate in Puerto Rico is
   a. 30 percent
   b. 62 percent
   c. 88 percent
5. A land reform program has been in operation for
   a. 10 years.
   b. 25 years.
   c. 40 years.

Inferential Questions (Understanding Ideas)

6. Puerto Rico came under the jurisdiction of the U.S.
   a. when it was purchased from a foreign country.
   b. when it was captured during a rebelling.
   c. when it was yielded to the U.S. as a war settlement.

7. A land reform program in Puerto Rico
   a. has changed the appearance of the waterfront
   b. has eliminated poverty
   c. has reduced American aid to the island.

8. The author states that Puerto Rico has
   a. a famous military museum.
   b. several colleges.
   c. beautiful gardens.

9. The author implies that the old part of San Juan
   a. has been rebuilt recently.
   b. is located on an island.
   c. is famous for its coral beds.

10. We can conclude that
    a. the economy of Puerto Rico is developing rapidly
    b. Puerto Rico imports large quantities of grains.
    c. most people cannot read or write in Puerto Rico.

Note: Correct answers are highlighted.

APPENDIX F

Reading Survey

Subject ID: ____________________

Please provide as accurate information as possible. The survey should not take more than three minutes.

DEMOGRAPHICS

Sex: ___Male___ Female **Age:** ____
Major:________________________________

Ethnicity: ________________________________ **Native Language:** ________________________

College standing: ___ Freshman___Sophomore___Junior___Senior

Did you transfer to GSU? ___Yes ___ No ___ N/A

Indicate the following:

I have taken the Regents' Reading Test / Compass Test ____ times, and my last score is _____.

What is your mother's educational level?

1- Less than high school
2- Some high school
3- High school graduate
4- Some college
5- College graduate

QUESTIONS ABOUT READING

1. How would you describe your reading ability?
    1. Can't read
    2. Poor reader
    3. Average reader

4. Better than average reader

2. Not including books required for school courses or your job, how many books do you typically read in a year?
    1. None
    2. 1-2
    3. 3-5
    4. 6-10

3. Which of the following is true?
    1. I have a library card for a community library.
    2. I do not have a community library card.

4. How many magazines do you yourself (not your family) subscribe to or purchase on a regular basis?
    1. None
    2. 1
    3. 2-5
    4. 6-10
    5. More than 10

5. I usually…
    1. Read more than one newspaper a day
    2. Read a newspaper everyday
    3. Read a daily newspaper occasionally
    4. Do not have time to read a daily newspaper
    5. Do not care to read a daily newspaper even if I had the time

6. How much television do you usually watch per day?
    1. I almost never watch television
    2. Less than one hour
    3. 1-3 hours
    4. 4-6 hours
    5. More than 6 hours

7. Approximately how often did your parents/guardians read to you when you were a child?
    1. None
    2. 1-2 times a week
    3. 3-4 times a week
    4. 5-6 times a week
    5. More than 6 times a week

8. How old were you when others in your family first began to read to you?
    1. Less than 6 months
    2. 6 months-1 year

3. 2 years-3 years
4. More than 3 years
5. I was never read to until I entered school

9. How often do you visit the school or public library?
    1. Never
    2. 1 time per month
    3. 2-4 times per month
    4. 5-10 times per month
    5. More than 10 times per month

10. Not including books for school or work, approximately how many books do you own?
    1. 0-10
    2. 11-20
    3. 21-30
    4. 31-50
    5. More than 50

11. How much time during the day, AT HOME, do you spend watching educational television programs and videos (e.g., Discovery Channel, National Geographic, etc.)?
    1. None
    2. 30 minutes
    3. 1 hour
    4. 2 hours
    5. 3 hours or more

12. How much time during the day, AT HOME, do you watch programs and videos for entertainment?
    1. None
    2. 30 minutes
    3. 1 hour
    4. 2 hours
    5. 3 hours or more

13. How often do you normally work at a job(s) a week?
    1. None
    2. 1-10 hours
    3. 11-20 hours
    4. 21-40 hours
    5. More than 40 hours

14. How would you rate your level of satisfaction with your job?
    1. I do not have a job
    2. Not satisfied at all
    3. Hardly satisfied

4. Average
5. Extremely satisfied

15. What is your current college enrollment status?
    1. Not enrolled
    2. Part time
    3. Full time

16. What is your current college G.P.A.?
    1. I don't know
    2. Less than 2.0
    3. 2.1-3.0
    4. 3.1-3.9
    5. 4.0

17. How well informed would you say you are about events happening in the world?
    1. Not informed at all
    2. Somewhat informed
    3. Average
    4. More than average
    5. Extremely well informed

18. What do you think your chances are of getting a good job after graduating from college?
    1. I don't know
    2. Somewhat good
    3. Average
    4. Better than average
    5. Extremely good

19. How would you rate the importance of reading to your academic success in college?
    1. I don't know
    2. Not important
    3. Somewhat important
    4. Average importance
    5. Extremely important

20. How many hours of study do you devote to your classes a week?
    1. None
    2. 1-2 hours
    3. 3-4 hours
    4. 5-6 hours
    5. More than 6 hours

21. How important to you is it to graduate from college?
    1. Not important

2. Somewhat important
3. Average importance
4. More than average importance
5. Extremely important

22. How important is it to you to continue your studies through graduate school?
    1. I do not plan on going to graduate school
    2. Somewhat important
    3. Average importance
    4. More than average importance
    5. Extremely important

23. How important is it to you to receive/maintain the HOPE scholarship in college?
    1. Not important
    2. Somewhat important
    3. Average importance
    4. More than average importance
    5. Extremely important
    6. Not Applicable

APPENDIX G

Instructions for the Error Detection Task

The following instructions appeared on the initial screen of the error detection task. Before starting the experiment, subjects were allowed to ask questions, and they were walked through a practice trial to familiarize them with the experiment.

Welcome to the story comprehension experiment.

In this experiment, you will be asked to read short stories and to answer some simple questions about them. The stories will be presented one line at a time on your computer screen. When you have read and comprehended a line of text, press the space bar and the next line will appear. Continue reading and pressing the space bar until you reach the end of each story. At that time, you will see the word QUESTION appear on the screen. This is a signal that a question about the story will soon appear. When the question appears, press the YES key if it is true about the story you just read, and press the NO key if it is not true about the story. Please keep your thumbs resting on the keyboard and your index fingers on the YES and NO keys.

It is important that you read at your normal pace and that you answer the comprehension questions as accurately as possible.

Let's do a practice.

Press the spacebar to begin.

APPENDIX H

Instructions for the Reading Span Task

The following instructions appeared on the initial screen of the task. Before starting the experiment, subjects were allowed to ask questions, and they were walked through a practice trial to familiarize them with the experiment.

Welcome to the experiment.

In this task, you will be presented with a series of unrelated sentences.
Whenever a sentence is presented to you, I want you to read each
sentence carefully. Some of the sentences make sense, and some
of the sentences don't make sense. After you finish reading each sentence, I
want you to click YES if the sentence makes sense and click NO if the sentence
doesn't make sense. When you're deciding whether or not the sentences make sense,
keep in mind that I'm not trying to trick you with hidden meanings or anything, so
don't waste too much mental energy over analyzing the sentences.

After clicking YES or NO, press the space bar and read the next sentence that is
presented, again clicking, YES or NO to indicate whether or not the sentence
makes sense. Keep doing this until you see a question mark. The question mark means that the trial is over, and you have to say back the last word in each of the sentences in the trial.

So, here's an example:

The astronaut placed a flag on the moon.
You will click YES....

The grass in the garden was pink.
You will click NO...

If possible, you are to say back the last words in the order in which they were
presented (moon, pink). If you can't remember them in order, you can say them in
any order, but you should not start with the last word first, unless it is the only one
you can remember. Your goal is to try to say back as many of the last words in
the trial as possible. We will be starting off with trials consisting of two sentences and
will periodically increase the number of sentences per trial without any advance warning.

That is, we will progress to three-sentence trials, then four-sentence trials and so on. The first couple of trials are for practice, so you can get the hang of it.

Let’s practice.
Press the spacebar to begin.

APPENDIX I

Author Recognition Test (Acheson et al., 2008)

Participant Number:______ Score: C_____I_____C-I_____

Below is a list of names. Some of them are authors of books, and some of them are not. Please put a check mark next to the ones that you know for sure are authors. There is a penalty for guessing, so you should check only those names about which you are absolutely certain. Thank you.

___Patrick Banville	___Harry Coltheart	___Virginia Woolf	___Tony Hillerman
___Kristen Steinke	___Gary Curwen	___John Landau	___Amy R. Baskin
___Ernest Hemingway	___Herman Wouk	___Toni Morrison	___James Clavell
___Clive Cussler	___Geoffrey Pritchett	___Harriet Troudeau	___Salmon Rushdie
___Hiroyuki Oshita	___Ray Bradbury	___Roswell Strong	___Maryann Phillips
___Kurt Vonnegut	___Jay Peter Holmes	___J.R.R. Tolkien	___Scott Alexander
___Anne McCaffrey	___Christina Johnson	___Margaret Atwood	___Ayn Rand
___Elinor Harring	___Jean M. Auel	___Seamus Huneven	___Alex D. Miles
___Sue Grafton	___Judith Stanley	___Harper Lee	___Margaret Mitchell
___Lisa Woodward	___Gloria McCumber	___Chris Schwartz	___Leslie Kraus
___David Harper Townsend	___James Joyce	___Walter LeMour	___Ralph Ellison
___Anna Tsing	___Robert Ludlum	___Alice Walker	___Sidney Sheldon
___T.C. Boyle	___Larry Applegate	___Elizabeth Engle	___ Brian Herbert
___Jonathan Kellerman	___Keith Cartwright	___T.S. Elliot	___Sue Hammond
___Cameron McGrath	___Jackie Collins	___Marvin Benoit	___Jared Gibbons
___F. Scott Fitzgerald	___Umberto Eco	___Joyce Carol Oates	___Michael Ondaatje
___A.C. Kelly	___David Ashley	___Jessica Ann Lewis	___Thomas Wolfe
___Peter Flaegerty	___Jack London	___Nelson Demille	___Jeremy Weissman
___Kazuo Ishiguro	___Seth Bakis	___Arturo Garcia Perez	___Willa Cather
___Jane Smiley	___Padraig O'seaghdha	___S.L. Holloway	___J.D. Salinger
___James Patterson	___E.B. White	___John Irving	___ Antonia Cialdini
___Martha Farah	___Giles Mallon	___Stephen Houston	___ Lisa Hong Chan
___Craig DeLord	___Raymond Chandler	___Marcus Lecherou	___Samuel Beckett
___Nora Ephron	___Isabel Allende	___Valerie Cooper	___Beatrice Dobkin
___Ann Beattie	___Amy Graham	___Tom Clancy	___Wally Lamb
___Stewart Simon	___Marion Coles Snow	___Vladimir Nabokov	___Katherine Kreutz
___Danielle Steel	___George Orwell	___Pamela Lovejoy	___James Michener
___Dick Francis	___Maya Angelou	___Vikram Roy	___William Faulkner
___Ted Mantel	___Bernard Malamud	___Saul Bellow	___Isaac Asimov
___I.K. Nachbar	___John Grisham	___Stephen King	___Lindsay Carter
___Judith Krantz	___Erich Fagles	___Elizabeth May Kenyon	___Paul Theroux
___Thomas Pynchon	___Walter Dorris	___Frederick Mundow	___Francine Preston
___Wayne Fillback	___Gabriel Garcia Marquez		

Please check the statement that best estimates how much time you spend reading in relation to other college students. In your estimate, remember to take into account all forms of reading, including magazines, newspapers, pleasure-reading, class assignments, email, etc.

_____ I read more often than the average college student.
_____ I read about as often as the average college student.
_____ I read less often than the average college student.

APPENDIX J

Metacognitive Awareness of Reading Strategies Inventory

Student ID: ______________

Metacognitive Awareness of Reading Strategies Inventory
(MARSI) Version 1.0
Kouider Mokhtari and Carla Reichard © 2002

DIRECTIONS: Listed below are statements about what people do when they read academic or school-related materials such as textbooks, library books, etc. Five numbers follow each statement (1, 2, 3, 4, 5) and each number means the following:

- **1** means "I **never or almost never** do this."
- **2** means "I do this **only occasionally.**"
- **3** means "I **sometimes** do this." (About **50%** of the time.)
- **4** means "I **usually** do this."
- **5** means "I **always or almost always** do this."

After reading each statement, **circle the number** (1, 2, 3, 4, or 5) that applies to you using the scale provided. Please note that there are **no right or wrong answers** to the statements in this inventory.

TYPE	STRATEGIES	SCALE				
GLOB	1. I have a purpose in mind when I read.	1	2	3	4	5
SUP	2. I take notes while reading to help me understand what I read.	1	2	3	4	5
GLOB	3. I think about what I know to help me understand what I read.	1	2	3	4	5
GLOB	4. I preview the text to see what it's about before reading it.	1	2	3	4	5
SUP	5. When text becomes difficult, I read aloud to help me understand what I read.	1	2	3	4	5
SUP	6.I summarize what I read to reflect on important information in the text.	1	2	3	4	5
GLOB	7. I think about whether the content of the text fits my reading purpose.	1	2	3	4	5
PROB	8. I read slowly but carefully to be sure I understand what I'm reading.	1	2	3	4	5
SUP	9.I discuss what I read with others to check my understanding.	1	2	3	4	5
GLOB	10. I skim the text first by noting characteristics like length and organization.	1	2	3	4	5
PROB	11. I try to get back on track when I lose concentration.	1	2	3	4	5
SUP	12. I underline or circle information in the text to help me remember it.	1	2	3	4	5
PROB	13. I adjust my reading speed according to what I'm reading.	1	2	3	4	5
GLOB	14. I decide what to read closely and what to ignore.	1	2	3	4	5
SUP	15. I use reference materials such as dictionaries to help me understand what I read.	1	2	3	4	5
PROB	16. When text becomes difficult, I pay closer attention to what I'm reading.	1	2	3	4	5
GLOB	17. I use tables, figures, and pictures in text to increase my understanding.	1	2	3	4	5
PROB	18. I stop from time to time and think about what I'm reading.	1	2	3	4	5
GLOB	19. I use context clues to help me better understand what I'm reading.	1	2	3	4	5
SUP	20. I paraphrase (restate ideas in my own words) to better understand what I read.	1	2	3	4	5
PROB	21. I try to picture or visualize information to help remember what I read.	1	2	3	4	5
GLOB	22. I use typographical aids like bold face and italics to identify key information.	1	2	3	4	5
GLOB	23. I critically analyze and evaluate the information presented in the text.	1	2	3	4	5
SUP	24. I go back and forth in the text to find relationships among ideas in it.	1	2	3	4	5
GLOB	25. I check my understanding when I come across conflicting information.	1	2	3	4	5
GLOB	26. I try to guess what the material is about when I read.	1	2	3	4	5
PROB	27. When text becomes difficult, I re-read to increase my understanding.	1	2	3	4	5
SUP	28. I ask myself questions I like to have answered in the text.	1	2	3	4	5
GLOB	29. I check to see if my guesses about the text are right or wrong.	1	2	3	4	5
PROB	30. I try to guess the meaning of unknown words or phrases.	1	2	3	4	5

Reference: Mokhtari, K., & Reichard, C. (2002). Assessing students' metacognitive awareness of reading strategies.

APPENDIX K

Directions for the Training Conditions

Repeated Readings

This binder was put together for the Repeated Readings condition. The experimenter keeps the binder and brings it to each session. Included in the binder are (a) *Timed Readings* passages, (b) the Answer Key, (c) the Progress Graph, (d) the Time-Rate Conversions Table, and (e) maze tests.

Follow the instructions listed below carefully:

1. Retrieve a pencil from the instructor.
2. Review your Intervention Timetable (next page).
3. Complete a maze test at sessions 4 and 7. A maze is also referred to as a Passage Completion Test. Take the maze before you start the training session.
4. Review your Progress Graph. See how you are doing on meeting the instructional challenge. The instructional challenge is defined in a given session as **reading a passage in 1 minute three times back to back and answering 8 questions correctly (80% accuracy)**. **If you meet this challenge in three sessions back to back, you will be promoted to the next reading level.**
5. Locate the passage to read this session.
6. Set your goal for this session as the instructional challenge: read the passage in **1 minute** with **80% accuracy (**a total of **8 questions correct).**
7. (**group/individual administration**) Begin reading when you are told by the instructor to do so.
8. (**group administration**) Look up to the large-screen timer when you are finished reading and copy the time you see on the screen on the upper right-hand corner of the passage.
9. (**individual administration**) Say "done!" when you finish reading the passage.
10. (**individual administration**) Listen to the instructor tell you your reading time.
11. (**group/individual administration**) Refer to the Time-Rate Conversions table to derive your words-per-minute rate using your reading time.
12. (**group/individual administration**) Plot your reading rate on the Progress Graph.

13. (**group/individual administration**) Repeat the procedures 5-12 *three more times*.
14. (**group/individual administration**) Answer the comprehension questions on the back of the page and mark your answers.
15. (**group/individual administration**) Refer to the answer key to check your answers.
16. (**group/individual administration**) Write down your comprehension score on the Progress Graph.
17. You are done!
18. Return your binder to the instructor.

Wide Reading

This binder was put together for the Wide Reading fluency group. The experimenter keeps the binder and brings it to each session. Included in the binder are (a) *Timed Readings* passages, (b) the Answer Key, (c) the Progress Graph, (d) the Time-Rate Conversions Table, and (e) maze tests.

Follow the instructions listed below carefully:

1. Retrieve a pencil from the instructor.
2. Review your Intervention Timetable (next page).
3. Complete a maze test at sessions 4 and 7. A maze is also referred to as a Passage Completion Test. Take the maze before you start the training session.
4. Review your Progress Graph. See how you are doing on meeting the instructional challenge. The instructional challenge is defined in a given session as **reading three consecutive passages in 1 minute and answering 10 questions correctly out of a total of 12 questions (80% accuracy)**. **If you meet this challenge in three sessions back to back, you will be promoted to the next level.**
5. Locate the 4 passages to read this session.
6. Set your goal for this session as the instructional challenge: reading at **400 wpm (1 minute)** on **three consecutive readings** with **80% accuracy (**a total of **10 questions correct).**
7. Turn to the *first passage* to read this session.
8. (**group/individual administration**) Begin reading the passage when you are told by the instructor to do so.
9. (**group administration**) Look up to the large-screen timer when you finish reading and copy the time you see on the upper right-hand corner of the passage.
10. (**individual administration**) Say “done!” when you finish reading the passage.

11. (**individual administration**) Listen to the instructor tell you your reading time.
12. (**group/individual administration**) Refer to the Time-Rate Conversions table to derive your words-per-minute rate using your reading time.
13. (**group/individual administration**) Plot your reading rate on the Progress Graph for the reading.
14. (**group/individual administration**) Answer questions 4, 5, and 6 on the back of the page and mark your answers.
15. (**group/individual administration**) Refer to the answer key to check your answers.
16. (**group/individual administration**) Write down your comprehension score on the Progress Graph.
17. (**group/individual administration**) Turn to the *second passage* to read for this session.
18. (**group/individual administration**) Repeat procedures 8-16 for the *second passage*.
19. (**group/individual administration**) Turn to the *third passage* to read for this session.
20. (**group/individual administration**) Repeat the procedures 8-16 for the *third passage*.
21. (**group/individual administration**) Turn to the *fourth passage* to read for this session.
22. (**group/individual administration**) Repeat the procedures 8-16 for the *fourth passage*.
23. You are done! Return your binder to the instructor.

Vocabulary Study

This binder was put together for Project RIFLE. The experimenter keeps the binder and brings it to each session. Included in the binder are (a) Vocabulary Study passages, (b) Vocabulary Builder Drill (c) the Answer Key, (c) the Progress Graph, (d) Vocabulary Study Cards, and (e) maze tests.

In this condition, Vocabulary Study, you will learn rare words by studying their spelling, pronunciation, and meaning.

Follow the instructions listed below carefully:

1. Retrieve a pencil from the instructor.
2. Review your Intervention Timetable (next page).
3. **(sessions 4 & 7)** Complete a maze test. A maze is also referred to as a Passage Completion Test. Take the maze before you start the training session.
4. Locate the Vocabulary Group to study for this session.

5. Study each word and its definition carefully.
6. Consult a dictionary if the definition provided is not clear or sufficient.
7. Take the Drill on the back of the page; mark and check your answers.
8. Write your score on the Intervention Timetable.
9. Identify 5 words that you missed or are not (very) familiar to you.
10. Fill out a Vocabulary Study Card for each of 5 words you selected for further study.
11. Write the word on the top of the Card indicating its part of speech (e.g., countable noun, uncountable noun, transitive verb, intransitive verb, adjective, adverb, pronoun, or preposition).
12. Look up the word in the dictionary.
13. Write down the word’s pronunciation from the dictionary entry.
14. Write down a context clue from the dictionary below the word’s pronunciation. You don't need to copy the whole sentence from the dictionary. Copy enough context to show the meaning of the word.
15. Write your own original example sentence using the word. Make sure your sentence is long enough for me to know that you understand what the word means. For example, it is difficult to say that you know the word “winter” in the sentence “I don’t like winter.” A sentence like “I don’t like winter because it is too cold and snowy most of the time” tells me more about how much you know about the word.
16. Write the word’s definition on the back of the Card
17. You are done!
18. Return your binder to the instructor.

APPENDIX L

Procedural Integrity Checklist

Repeated Readings

A binder will be put together for each Project RIFLE participant. The experimenter keeps the binder and brings it to each session. Included in the binder are (a) passages from the target *Timed Readings* level, (b) the Answer Key, (c) the Progress Graph, and (d) the Time-Rate Conversions table.

In a typical session, the procedures listed below are to be followed as closely as possible.

Before the session:

1. Check each subject's binder.
2. Review the records from the previous session (Progress Graph, Intervention Timetable, etc.). Recall the instructional criteria: achieving a **400-wpm (**in **1 minute) reading rate** on **three consecutive readings** with **80% accuracy (**a total of **8 questions correct** out of 10**) in 3 consecutive sessions**. When these criteria are met, students will be promoted to the next level.
3. Check in the Intervention Timetable to determine if the participant is taking a maze test this session. A maze test must be completed at Sessions 4 and 7.
4. Check in the Intervention Timetable to determine if the participant met the instructional criteria during the previous session(s).
5. If the subject read at the instructional criteria, include the first next-level passage.
6. If not, use the next same-level passage.
7. Copy the passages from the target series using copy ratio 125.
8. Copy the passage and questions double-sided.
9. Date the passage.
10. Date the session on the Progress Graph.
11. Update the Intervention Timetable.
12. Obtain a stopwatch, a pencil per subject.

13. Obtain a large screen stopwatch (i.e., on a laptop computer) if training multiple subjects.
14. Arrive early in the room.
15. Greet the students.

During the session:

16. (**group/individual administration**) Hand out the binders along with a pencil.
17. (**group/individual administration**) Ask students to review the instructions for their condition, Intervention Timetable, and their Progress Graph.
18. (**group/individual administration**) Ask students to locate the passage they will read this session.
19. (**group/individual administration**) Read aloud the instructional criteria to the students:

    > "If you read the passage for this session in **1 minute three times back to back and answer 8 questions (out of a total of 10) correctly**, you meet the challenge for this session. **If you meet the challenge in three sessions back to back, you will be promoted to the next reading level."**

20. (**group/individual administration**) Ask students to turn to the passage they will read this session.
21. (**group administration**) Point to the large screen stopwatch on a computer screen. Ask students to look up to the screen when they finish reading and to copy the time on the top right-hand corner of the passage.
22. (**individual administration**) Ask student to say "done!" when s/he finishes reading the passage.
23. (**group/individual administration)** Direct student(s) to begin reading silently. Start timing when they begin**.**
24. (**group administration**) Make sure the students are following the instructions. After the first reading, go around the room and check if they did copy their reading times from the large-screen timer on their passages.
25. (**group administration**) Stop the large screen timer when all the students finish reading.
26. (**individual administration**) Stop timing; tell the student her/his reading time.
27. (**group/individual administration**) Instruct students to refer to the Time-Rate Conversions table to derive their words-per-minute rate using their reading time from the reading.
28. (**group/individual administration**) Instruct students to plot their reading rate on the Progress Graph.

29. (**group/individual administration**) Repeat the procedures 21-28 *three more times* on the *same passage*.
30. (**group/individual administration**) Instruct students to answer all 10 comprehension questions on the back of the page and mark their answers.
31. (**group/individual administration**) Instruct students to refer to the answer key to check their answers.
32. (**group/individual administration**) Instruct students to write down their comprehension scores on the Progress Graph.
33. (**group/individual administration**) Review the Progress Graph with the subject(s).
34. (**group/individual administration**) Thank the student(s) for attending the session.
35. (**group/individual administration**) Collect the binders.
36. (**group/individual administration**) Dismiss the session.

After the session:

37. Indicate on the Intervention Timetable form if the subject reached the instructional criteria.
38. Write down observations from the session.

Wide Reading

A binder will be put together for each Project RIFLE participant. The experimenter keeps the binder and brings it to each session. Included in the binder are (a) passages from the target *Timed Readings* level, (b) the Answer Key, (c) the Progress Graph, and (d) the Time-Rate Conversions table.

In a typical session, the procedures listed below are to be followed as closely as possible.

Before the session:

1. Check each student's binder.
2. Review the records from the previous session (Progress Graph, Intervention Timetable, etc.). Recall the instructional criteria: achieving a **400-wpm (**in **1 minute) reading rate** on **three consecutive readings** with **80% accuracy (**a total of **10 questions correct** out of 12**) in 3 consecutive sessions.** When this goal is reached, subjects will be promoted to the next level.
3. Check in the Intervention Timetable to determine if the participant is taking a maze test this session. A maze test must be completed at Sessions 4 and 7.
4. Check in the Intervention Timetable to determine if the participant met the instructional criteria during the previous session(s).

5. If the instructional criteria were met, include in the binder the first four next-level passages.
6. If not, use the next four same-level passages.
7. Copy the passages from the target series using copy ratio 125.
8. Copy the passages and questions double-sided.
9. Date the passages.
10. Date the session on the Progress Graph.
11. Update the Intervention Timetable.
12. Obtain a stopwatch, a pencil per student.
13. Obtain a large screen timer (i.e., on a laptop computer) if training multiple subjects.
14. Arrive early in the room.
15. Greet the students.

During the session:

16. (**group/individual administration**) Hand out the binders along with a pencil.
17. (**group/individual administration**) Ask students to review the instructions for their condition, the Intervention Timetable, and their Progress Graph.
18. (**group/individual administration**) Ask students to locate the passages they will read this session.
19. (**group/individual administration**) Read aloud the instructional criteria to students:

    "You will read 4 passages today's training. If you finish reading in **1 minute on three consecutive passages answer 10 questions (out of a total of 12) correctly**, **you meet the challenge for this session**. **If you meet the challenge in three sessions back to back, you will be promoted to the next reading level."**

20. (**group/individual administration**) Ask students to turn to the *first passage* they will read this session.
21. (**group administration**) Point to the large screen stopwatch on a computer screen. Ask students to look up to the screen when they finish reading and to copy the time on the top right-hand corner of the passage.
22. (**individual administration**) Ask the student to say "done!" when s/he finishes reading the passage.
23. (**group/individual administration)** Direct student(s) to begin reading. Start timing when they begin**.**
24. (**group administration**) Stop the large screen stopwatch when all the students finish reading and copying their reading times on their passages.

25. (**group administration**) Make sure the students are following the instructions. After the first reading, go around the room and check if they did copy their reading times from the large-screen timer on their passages.
26. (**individual administration**) Stop timing; tell the student her/his reading time.
27. (**group/individual administration**) Instruct students to refer to the Time-Rate Conversions table to derive their words-per-minute rate using their reading time from the reading.
28. (**group/individual administration**) Instruct the students to plot their reading rate on the Progress Graph.
29. (**group/individual administration**) Instruct students to answer questions 4, 5, and 6 on the back of the page and mark their answers.
30. (**group/individual administration**) Instruct students to refer to the answer key to check their answers for this passage.
31. (**group/individual administration**) Instruct students to write down their comprehension scores on the Progress Graph for this passage.
32. (**group/individual administration**) Ask students to turn to the *second passage* they will read this session.
33. (**group/individual administration**) Repeat the procedures 21-31 for the *second passage*.
34. (**group/individual administration**) Ask students to turn to the *third passage* they will read this session.
35. (**group/individual administration**) Repeat the procedures 21-31 for the *third passage*.
36. (**group/individual administration**) Ask students to turn to the *fourth passage* they will read this session.
37. (**group/individual administration**) Repeat the procedures 21-31 for the *fourth passage*.
38. (**group/individual administration**) Review the Progress Graph with the student(s).
39. (**group/individual administration**) Thank the student(s) for attending the session.
40. (**group/individual administration**) Collect the binders.
41. (**group/individual administration**) Dismiss the session.

After the session:

42. Indicate on the Intervention Timetable form if the subject reached the instructional criteria.
43. Write down observations from the session.

Vocabulary Study

This binder was put together for Project RIFLE. The experimenter keeps the binder and brings it to each session. Included in the binder are (a) Vocabulary Study passages, (b) Vocabulary Builder Drill (c) the Answer Key, (c) the Progress Graph, (d) Vocabulary Study Cards, and (e) maze tests.

In this condition, Vocabulary Study, students will learn rare words by studying their spelling, pronunciation, and meaning.

In a typical session, the procedures listed below are to be followed as closely as possible.

Before the session:

1. Check each student's binder.
2. Review the records from the previous session (i.e., Intervention Timetable).
3. **(sessions 2 and on)** Review the Vocabulary Study Cards written by students. Check if all components of the assignment were completed; make sure an original sentence was written for each Word Card. Read the students' original sentences and give suggestions as to how the sentence may be revised if the sentences are too short or do not reflect knowledge of the words.
4. Check in the Intervention Timetable to determine if the participant is taking a maze test this session. A maze test must be completed at Sessions 4 and 7.
5. Include in the binder the next Vocabulary Study Group words.
6. Copy the passages from the target series using copy ratio 125.
7. Copy the passages and questions double-sided.
8. Date the passages.
9. Date the session on the Progress Graph.
10. Update the Intervention Timetable.
11. Obtain a pencil per student and dictionary(ies).
12. Arrive early in the room.
13. Greet the students.

During the session:

1. (**group/individual administration**) Hand out the binders along with a pencil.
2. (**group/individual administration**) Ask students to review the instructions for their condition and the Intervention Timetable.
3. (**group/individual administration**) Ask students to locate the Vocabulary Study Group they will complete this session.
4. **(sessions 4 & 7)** (**group/individual administration**) Have students complete a maze test. A maze is also referred to as a Passage Completion Test. Administer the maze prior to the training session.

5. **(group/individual administration)** Ask students to locate the Vocabulary Group to study for this session.
6. **(group/individual administration)** Ask students to study each word and its definition carefully.
7. **(group/individual administration)** Ask students to consult a dictionary if the definition provided is not clear or sufficient.
8. **(group/individual administration)** Ask students to take the Drill on the back of the page and to mark and check their answers.
9. **(group/individual administration)** Ask students to write their score on the Intervention Timetable.
10. **(group/individual administration)** Ask students to identify 5 words that they missed on the quiz or are not (very) familiar to them.
11. **(group/individual administration)** Ask students to fill out a Vocabulary Study Card for each of 5 words they selected for further study.
12. **(group/individual administration)** Ask students to write the word on the top of the Card indicating its part of speech (e.g., countable noun, uncountable noun, transitive verb, intransitive verb, adjective, adverb, pronoun, or preposition).
13. **(group/individual administration)** Ask students to look up the word in the dictionary.
14. **(group/individual administration)** Ask students to write down the word's pronunciation from the dictionary entry.
15. **(group/individual administration)** Ask students to write down a context clue from the dictionary below the word's pronunciation. Tell them they don't need to copy the whole sentence from the dictionary; they are to write down sufficient context for the word under study.
16. **(group/individual administration)** Ask students to write their own original example sentence using the word. The original sentence is to be long enough indicate that the students understand the word's meaning. For example, it is difficult to say that a student knows the word "winter" in the sentence "I don't like winter." A sentence like "I don't like winter because it is too cold and snowy most of the time" is more informative about how much the student knows about the word.
17. **(group/individual administration)** Ask students to write the word's definition on the back of the Word Study Card.
18. (**group/individual administration**) Review the Word Study Cards with the student(s).
19. (**group/individual administration**) Thank the student(s) for attending the session.
20. (**group/individual administration**) Collect the binders.
21. (**group/individual administration**) Dismiss the session.

After the session:

22. Write down observations from the session.

APPENDIX M

Descriptive and Normality Statistics of the Dependent Variables at Pretest and Posttest

	Pretest		Posttest	
	Mean (*SD*) [range]	Normality Test (Shapiro-Wilk)	*Mean* (*SD*) [range]	Normality Test (Shapiro-Wilk)
ND Reading Comprehension (maximum score=76)				
Wide Reading	34 (1.048) [22,48]	*W*= .878, *df*= 9, *p*=.150	34.44 (1.033) [20,50]	*W*= .940, *df*= 9, *p*= .577
Repeated Readings	31.27 (1.221) [12,48]	*W*= .929, *df*= 11, *p*=.398	28.73 (8.545)[14,40]	*W*= .935, *df*= 11, *p*= .465
Vocabulary Study	36 (9.285) [22, 48]	*W*=.932, *df*= 10, *p*= .467	29 (1.151)[14, 46]	*W*= .895, *df*= 10, *p*= .192
Comparison Group	60 (6.683)[52, 70]	*W*= .879, *df*= 13, *p*=.068		
ND Vocabulary (maximum score=80)				
Wide Reading	37.67 (7.549) [29, 51]	*W*= .981, *df*= 9, *p*= .105	43.22 (9.641)[32, 61]	*W*= .888, *df*= 9, *p*= .191
Repeated Readings	38.55 (7.448)[27.00, 53]	*W*= .981, *df*= 11, *p*= .970	42.27 (10.189)[29, 57]	*W*= .900, *df*= 11, *p*= .187
Vocabulary Study	33.40 (7.531)[25, 49]	*W*= .880, *df*= 10, *p*= .130	40 (6.815)[26, 47]	*W*= .871, *df*= 10, *p*= .103
Comparison Group	55.85 (12.456)[36, 74]	*W*=.922, 13, *p*= .265		
ND Silent Reading Rate				
Wide Reading	193.44 (49.709) [139, 305]	*W*= .848, *df*= 9, *p*= .071	249.44 (77.408) [179, 398]	*W*= .834, *df*= 9, *p*= .049*
Repeated Readings	215.45 (120.89) [116, 566]	*W*= .601, *df*= 11, p< .001*	226.27 (48.271) [137, 325]	*W*= .966, *df*= 11, *p*= .839
Vocabulary Study	182.30 (41.150) [116, 233]	*W*= .882, *df*= 10, *p*= .139	203.10 (70.291) [98, 299]	*W*= .907, *df*= 10, *p*= .262
Comparison Group	275.62 (63.69) [196, 388]	*W*= .923, *df*= 13, *p*= .274		
MARSI+ (maximum score=5)				
Wide Reading	3.48 (.601) [2.63,4.67]	*W*= .956, *df*= 9, *p*=.754	3.67 (.592) [3.10, 4.90]	*W*= .859, *df*= 9, *p*= .094
Repeated Readings	3.55 (.718) [2.30,4.57]	*W*= .968, *df*= 11, *p*= .869	3.76 (.436) [3.10,4.40]	*W*= .956, *df*= 11, p =.722
Vocabulary Study	3.10 (.769) [1.83,4.17]	*W*= .951, *df*= 10, *p*= .685	3.49 (.647) [2.60,4.50]	*W*= .951, *df*= 10, *p*= .679
Comparison Group	3.55 (.718) [2.30, 4.57]	*W*= .955, *df*= 13, *p*= .671		
ART (maximum score=65)				
Wide Reading	3.67 (3.162) [0, 9]	*W*= .936, *df*= 9, *p*= .542		
Repeated Readings	3.27 (2.723) [-1,10]	*W*= .855, *df*= 11, *p*= .049*		
Vocabulary Study	4.3 (3.401) [0, 8]	*W*= .846, *df*= 10, p = .052		
Comparison Group	12.23 (5.974) [4,22]	*W*= .925, *df*= 13, *p*= .292		
Maze (maximum score $_{pretest}$= 49; maximum score $_{posttest}$=56)				
Wide Reading	25.89 (9.584) [8,38]	*W*= .954, *df*= 9, *p*= .733	31.22 (1.151) [16,51]	*W*= .968, *df*= 9, *p*= .876
Repeated Readings	20.91 (5.839) [15,33]	*W*=.882, *df*= 11, *p*= .110	25.18 (6.705) [14,34]	*W*= .943, *df*=11, *p*= .560
Vocabulary Study	19.5 (4.035) [14, 25]	*W*= .912, *df*= 10, *p*= .297	25.8 (8.011) [16, 36]	*W*= .882, *df*= 10, *p*= .138
Comparison Group	32.85 (7.022) [21,41]	*W*= .852, *df*= 15, *p*= .031*		
Reading Span (maximum score= 70)				
Wide Reading	42.22 (13.017) [18,59]	*W*=.937, *df*= 9, *p*= .548	53.75 (6.541) [44,64]	*W*= .973, *df*= 9, *p*= .916
Repeated Readings	38.11 (13.986) [24,62]	*W*= .907, *df*= 11, *p*= .224	39.44 (17.133) [19,66]	*W*= .947, *df*= 11, *p*= .602
Vocabulary Study	47 (10.863) [30,63]	*W*= .966, *df*= 10, *p*= .856	50.33 (14.663) [19,65]	*W*= .827, *df*= 10, *p*= .031*
Comparison Group	52.62 (12.285) [32,69]	*W*= .948, *df*= 13, *p*= .570		

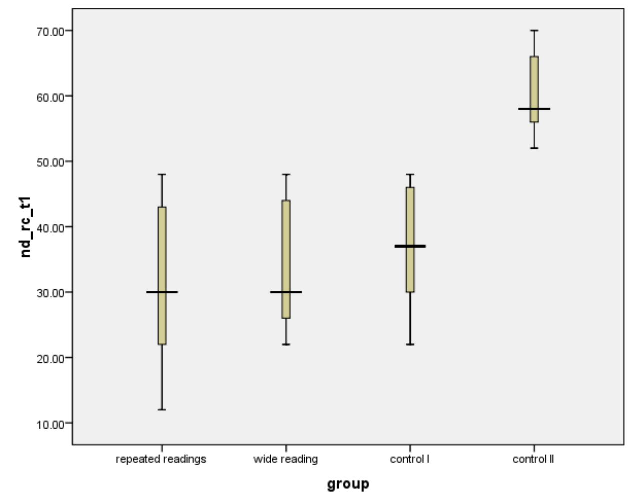
TOWRE SWE (maximum score=104)				
Wide Reading	88.46 (9.287) [68,94]	*W*=.931, *df*= 9, *p*= .487	81.33(10.05) [70,96]	*W*= .887, *df*= 9, *p*= .184
Repeated Readings	81.4 (11.423) [64,102]	*W*= .949, *df*= 11, *p*= .632	78.8 (12.173) [59,102]	*W*= .950, *df*= 11, *p*= .650
Vocabulary Study	74.67 (7.416) [63,84]	*W*= .958, *df*= 10, *p*= .763	72.33(6.364) [65,84]	*W*= .932, *df*= 10, *p*= .468
Comparison Group	88.27 (9.597) [70,104]	*W*= .966, *df*= 13, *p*= .836		
TOWRE PDE (maximum score= 63)				
Wide Reading	27.89 (11.494) [13, 42]	*W*= .871, *df*= 9, *p*= .126	28.11 (13.679) [9,47]	*W*= .941, *df*= 9, *p*= .597
Repeated Readings	27.4 (13.335) [9,47]	*W*=.945, *df*= 11, *p*= .575	24 (12.329) [9,46]	*W*= .928, *df*= 11, *p*= .395
Vocabulary Study	25.78 (10.86) [11,41]	*W*= .950, *df*= 10, *p*= .674	25.56 (11.204) [12,43]	*W*= .926, *df*= 10, *p*= .412
Comparison Group	38.31 (14.756) [9,59]	*W*= .963, *df*= 13, *p*= .795		

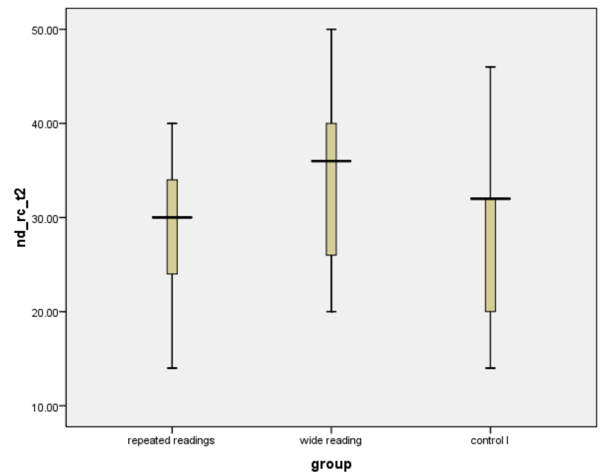
$^{+}$Judgments were made on a 5-point scale (1 = I never or almost never do this, 5 = I always or almost always do this).

APPENDIX N

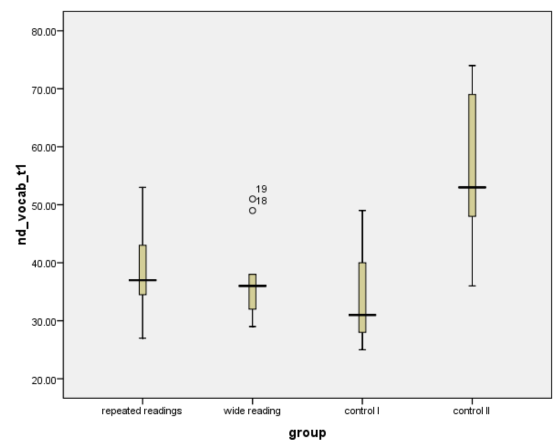
Box Plots for DVs at Pretest and Posttest

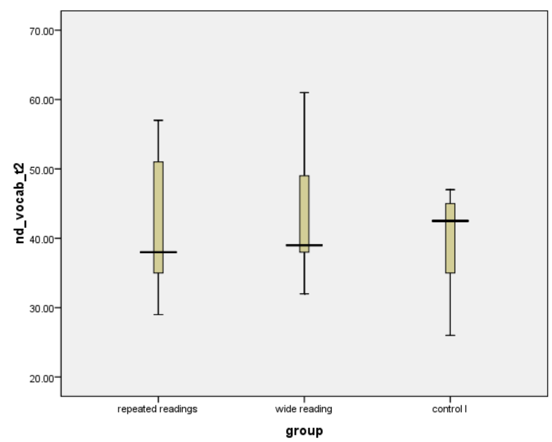
Pretest	Posttest



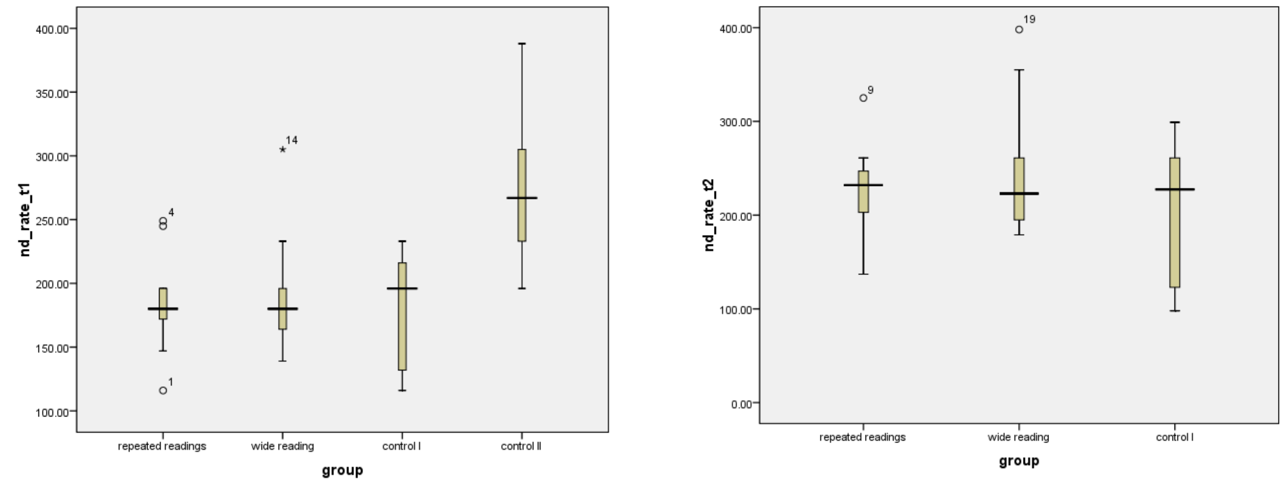


ND Reading Comprehension raw scores by group

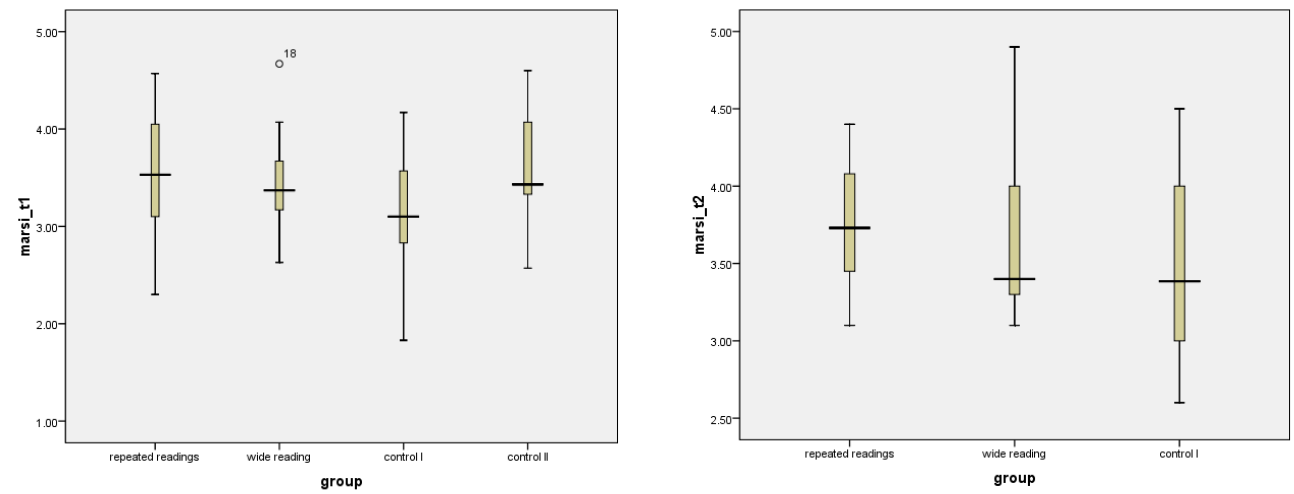




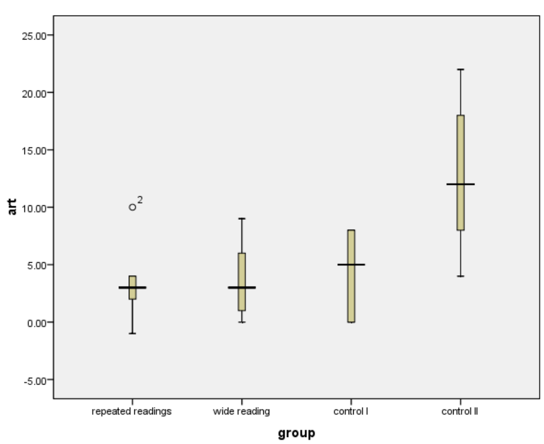
ND Vocabulary raw scores by group



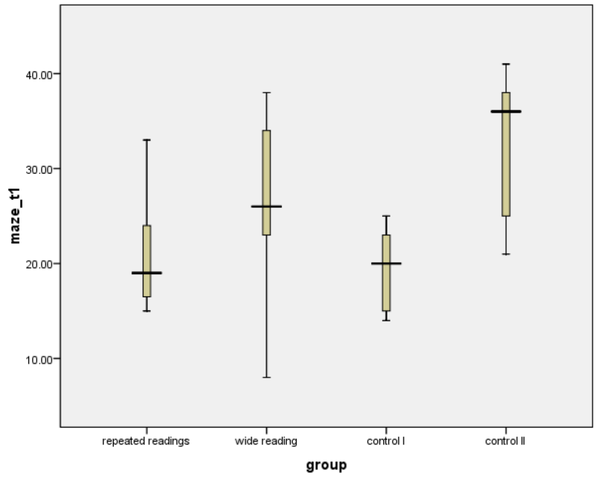
ND Silent Reading Rate raw scores by group

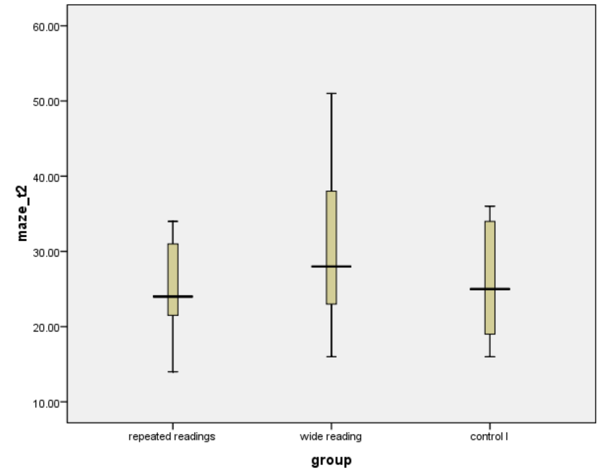


MARSI raw scores by group

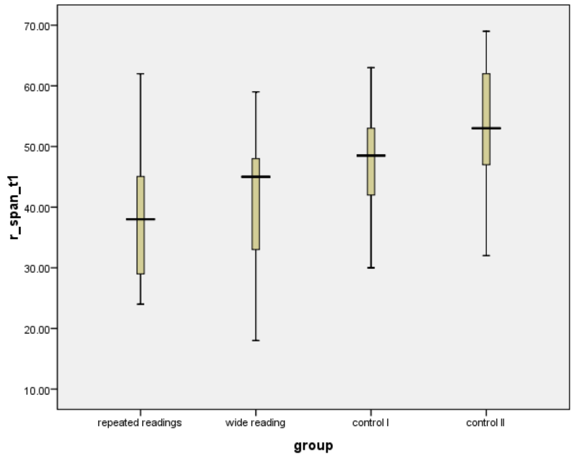


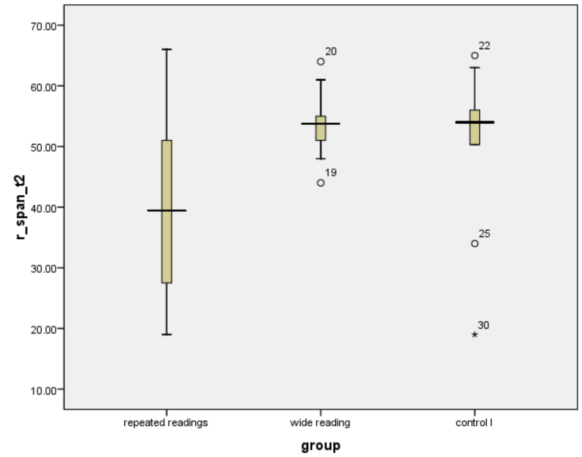
ART raw scores by group



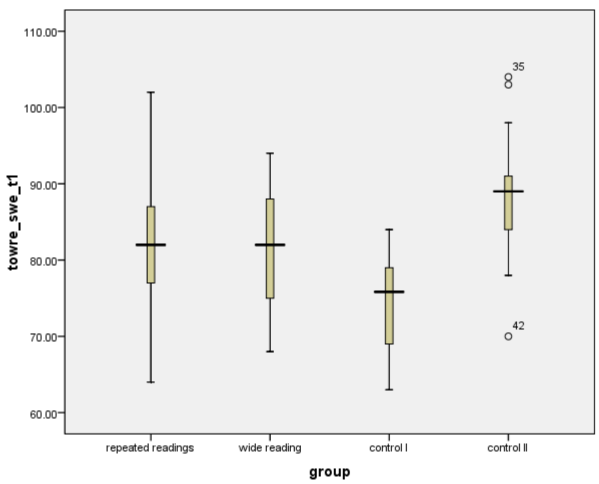


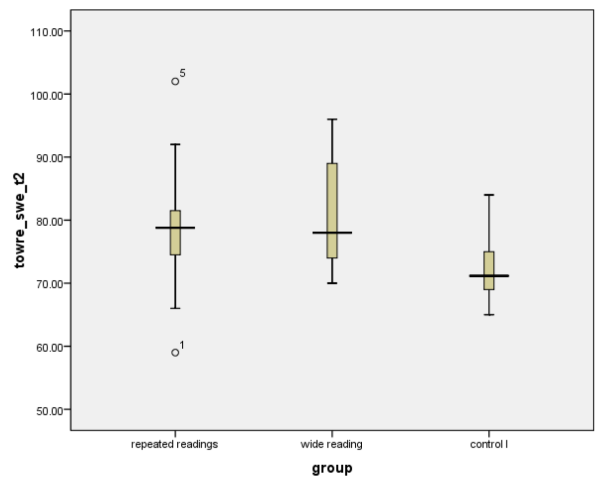
Maze raw scores by group



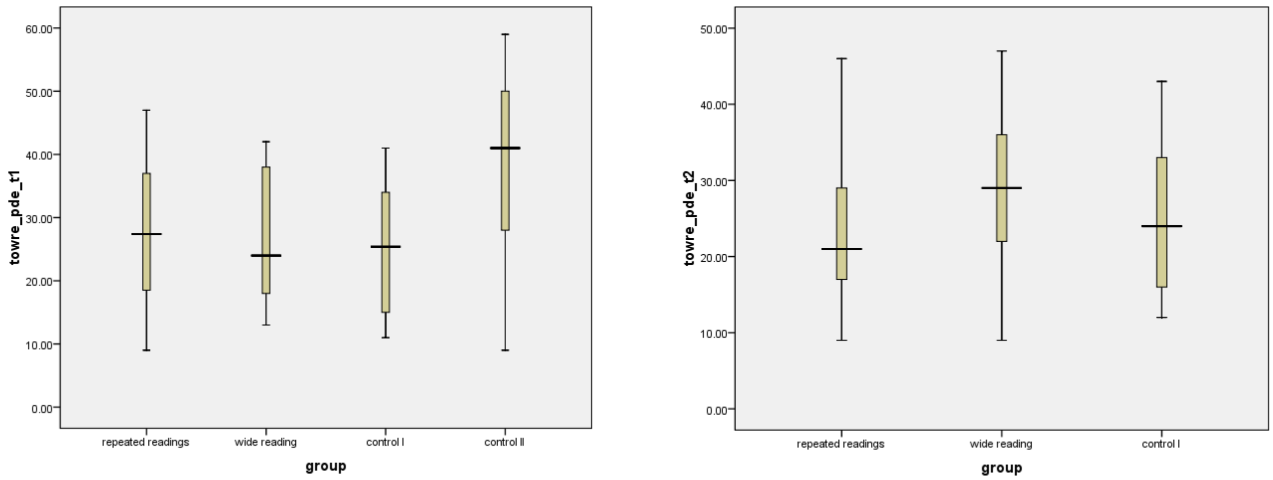


RSPAN Raw Scores by Group





TOWRE SWE raw scores by group



TOWRE PDE raw scores by group