

A META-ANALYSIS OF THE EFFECTS OF INTERVENTIONS TO INCREASE  
READING FLUENCY AMONG ELEMENTARY SCHOOL STUDENTS

By

Jiaxiu Yang

Dissertation

Submitted to the Faculty of the  
Graduate School of Vanderbilt University  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

in

Education and Human Development

August, 2006

Nashville, Tennessee

Approved:

Professor Douglas Fuchs

Professor Mark Lipsey

Professor Lynn Fuchs

Professor Daniel Reschly

Professor Donald Compton

## ACKNOWLEDGEMENTS

First and foremost I thank my advisor and chairperson, Professor Douglas Fuchs, for his guidance and unflagging support that made this work possible. Mentoring me through every step, he taught me more than I could have asked and in ways that made it directly relevant.

I am especially indebted to my minor advisor, Professor Mark Lipsey, who provided invaluable guidance and timely support throughout the data analysis decisions. Through his sharp technical perspectives, attention to detail, understanding, and persistent encouragement when the next steps seemed insurmountable, Professor Lipsey helped me complete this work.

I also thank my other committee members, Professor Lynn Fuchs, Professor Daniel Reschly, and Professor Donald Compton. Each offered me extensive academic and professional guidance and taught me a great deal about scientific research and life in general.

Additionally, I thank RMC Research Corporation at Portsmouth, NH, my employer, for being so incredibly supportive and understanding of my efforts to complete this project.

Nobody has been more important to me in the pursuit of this project than the members of my family. I would like to thank my parents, whose love, dedication, and sacrifice made it possible for me to pursue my dreams. Although my father unexpectedly died months before I completed this project, I take solace in knowing that he loved me and was proud of my work. I hope that he also knew how much I loved him and valued him as a role model for my own life. Finally, I give unending thanks to my loving and supportive husband, Jim, most dear to my heart, and to my wonderful daughter, Alana, who continues to provide inspiration in my life.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS .....	ii
LIST OF TABLES.....	vi
Chapter	
I. STATEMENT OF THE PROBLEM.....	1
Reading Problems are Pervasive and Persistent.....	1
Factors Contributing to Pervasive and Persistent Reading Problems.....	3
Understanding Reading Fluency .....	5
Need to Synthesize Instructional Effectiveness .....	8
II. LITERATURE REVIEW.....	13
Theories about Reading Fluency Development .....	13
Resource-Based Theories of Reading Fluency.....	14
Interactive-Compensatory Models of Reading Fluency .....	15
Instance Theory/Information Encapsulation of Reading Fluency .....	16
A Review of Primary Studies Conducted to Improve Reading Fluency .....	18
Repeated Reading .....	19
Single Case/Subject Studies .....	19
Group Design Studies without a Control Group.....	22
One-Group Pretest Posttest Studies.....	23
Multiple Group No Control Studies.....	24
Controlled Studies .....	29
Repeated Measures Studies.....	29
Controlled Group Design Studies .....	31
Classwide Peer Tutoring/Reading PALS.....	34
Previewing .....	36
Neurological Impress Method (NIM).....	37
Oral Recitation Lesson .....	38
Computer-Assisted Instruction .....	39
Decoding Training .....	40
Specific Instructional Methods.....	41
Summary of Interventions.....	43
Treatment Effects Do Not Transfer.....	44
Methodological Limitations .....	46
III. METHODS .....	48
Eligibility Criteria for Study Inclusion .....	48
Retrieval of Research Reports .....	51
Coding of Research Reports .....	52
Effect Size Coding .....	53
Coding of Study Descriptors.....	55

Methodological Variables.....	55
Study Context .....	55
Treatments.....	55
Participants .....	56
Training Materials .....	57
Coding Reliability.....	57
Data Analysis .....	58
IV. RESULTS .....	61
Descriptive Analysis .....	62
Effect Size Adjustments .....	70
Adjustments of All Effect Sizes.....	70
Independence among Effect Sizes from Multiple Treatment Studies.....	73
A Summary of Effect Sizes for Outcomes .....	77
Overall Mean Effect Sizes for Outcomes.....	77
Mean Effect Sizes by Student Type .....	80
Mean Effect Sizes by Treatment Type .....	81
Recoding of Treatment Type .....	82
Mean Effect Sizes by Treatment Type.....	83
Mean Effect Sizes by Treatment Features .....	90
Treatment Duration.....	91
Treatment Frequency .....	92
Treatment Format.....	93
Treatment Focus.....	95
Treatment Session Length.....	96
Treatment Intensity .....	97
Total Treatment Time .....	98
Multiple Regression on the Variability in the Effect Sizes .....	99
Predictor Variables .....	99
Subject Characteristics Predictors.....	99
Treatment Characteristics Predictors .....	100
Treatment Type Predictors .....	101
Multiple Regression Results for Effect Sizes of Reading Rate.....	108
Multiple Regression Results for Effect Sizes of Reading Accuracy.....	108
Multiple Regression Results for Effect Sizes of Comprehension .....	109
V. DISCUSSION AND CONCLUSIONS .....	111
Limitations .....	111
Summary and Discussion of Findings .....	114
Overall Effects of fluency Building Strategies.....	114
Factors Affecting Effects of Treatments .....	116
Implications for Practice, Fluency Theories, and Research .....	120
Implications for Reading Practice .....	120
Repetitive Practice vs. Non-Repetitive Practice.....	120
Focus on Fluency and Comprehension .....	121
Differential Treatment Effects for Students with Disabilities .....	122
Implications for Fluency Theories.....	123
Implications for Future Research.....	125

Appendix

A. STUDY INCLUSION CRITERIA.....	127
B. STUDY LEVEL CODING MANUAL.....	131
C. EFFECT SIZE CODING MANUAL.....	138
D. STUDY LEVEL CODING FORM .....	146
E. EFFECT SIZE LEVEL CODING FORM .....	159
F. EXCLUDED STUDIES .....	168
REFERENCES.....	177

## LIST OF TABLES

Table	Page
1. Number of Samples-Single Treatment Studies.....	64
2. Number of Treatments and Samples-Multiple Treatment Studies .....	66
3. Treatment Characteristics .....	67
4. Subject Characteristics .....	68
5. Effect Size Distribution in Study Reports .....	69
6. Mean Effect Sizes for Outcomes .....	78
7. Mean Effect Sizes by Student Types .....	81
8.1. Mean Effect Sizes by Different Fluency Strategies (1).....	85
8.2. Mean Effect Sizes by Different Fluency Strategies (2).....	86
8.3. Mean Effect Sizes by Different Fluency Strategies (3).....	86
9.1. Mean Effect Size by All Different Reading Practice Levels.....	88
9.2. Mean Effect Sizes by Repetitive Reading vs. Non-Repetitive Reading .....	89
10. Mean Effect Sizes by Different Text Levels .....	90
11. Mean Effect Sizes by Treatment Duration .....	92
12. Mean Effect Sizes by Treatment Frequency .....	93
13. Mean Effect Sizes by Treatment Format.....	94
14. Mean Effect Sizes by Treatment Focus .....	96
15. Mean Effect Sizes by Treatment Session Length .....	97
16. Mean Effect Sizes by Treatment Intensity.....	98
17. Mean Effect Sizes by Total Treatment Time .....	99
18. Correlations between ES and predictors.....	105
19. Correlations between predictors .....	106
20. Multiple Regression Results.....	110

## CHAPTER I

### STATEMENT OF THE PROBLEM

#### Reading Problems are Pervasive and Persistent

Reading is necessary for a child's success in school and life (Anderson, Hiebert, Scott, & Wilkerson, 1985), thus, it is unfortunate that reading difficulties are among the most common problems experienced by school-aged children. A study conducted by the National Assessment of Educational Progress (NAEP) found that 44% of students in the 4<sup>th</sup>-grade were dysfluent on grade-level stories. This study also found that students low in fluency were also low in reading comprehension (Pinnell, Pikulski, Wixson, Campbell, Gough, & Beatty, 1995). According to the latest report released by NAEP on the reading performance of the nation's 4<sup>th</sup>-graders, in 2000, only 32% of the students performed at or above a "proficient" level, the level identified by the National Assessment Governing Board (NAGB) as the level that all students should reach (Donahue, Finnegan, Lutkus, Allen, & Campbell, 2001).

Students who experience reading difficulties often qualify for Chapter 1 or special education services and join a burgeoning population of students labeled learning disabled, highlighting the magnitude of reading problems in the nation (Allington & McGill-Franzen, 1989). This finding is indeed disturbing when we consider that over 5,000,000 children in the United States (1 out of 9) are currently participating in Chapter I services (The Commission on Chapter I, 1993), and more than 5,540,000 received special education services in the 1998-99 school year

(Twenty-second Annual Report to Congress on the Implementation of IDEA, 2000). Over half (50.8%) of the students with disabilities served under IDEA in 1998-99 were categorized as having specific learning disabilities. Among these children, a majority experiences problems with reading (Fuchs, Fuchs, Mathes, Lipsey, & Roberts, 2001).

Reading problems are not just pervasive but also persistent. In fact, growing evidence suggests that the reading problems experienced by low-performing readers worsen over time (e.g., Stanovich, 1986) and persist throughout the entire schooling process (Juel, 1988; McGill-frazen & Allington, 1991). For students who are behind in reading and literacy development, the opportunities to advance or catch up diminish over time (Good, Simmons, & Smith, 1998). Therefore, if children begin their school career as poor readers, they are likely to remain poor readers without appropriate remediation, and make little progress in other language art skills such as listening comprehension or writing (Juel, 1988). Furthermore, their reading failure sours their attitude toward reading, which in turn decreases their interest in and consequently their exposure to books and information, widening the gap between these children and their peers. By the middle of third grade, the transition from “learning to read” to “reading to learn” is well under way. Students who have not mastered the requirements of reading by the end of the third grade are not likely to gain the opportunities to further enhance their literacy skills. And the NAEP’s latest report on the reading performance of the nation’s 4<sup>th</sup>-graders, which shows that, in 2000, lower-performing students have lost ground as compared to their



performance in previous years, will attest to that (Donahue, Finnegan, Lutkus, Allen, & Campbell, 2001).

### Factors Contributing to Pervasive and Persistent Reading Problems

The pervasiveness and persistence of reading problems underscore the importance of understanding why many children are not learning to read satisfactorily. Over the years, researchers have tried to uncover the factors that contribute to the development of reading problems. Descriptive research on successful and unsuccessful readers completed by Golinkoff (1975-1976) recognized three factors: (a) speed of word recognition, (b) type of reading errors made, and (c) decoding unit size. With respect to word recognition speed, poor readers tend to read more slowly than good readers (Biemiller, 1977-1978; Cromer, 1970; Colinkoff & Rosinski, 1976; Katz & Wicklund, 1971; Pace & Golinkoff, 1976; Perfetti & Hogaboam, 1975). Poor readers tend to make more meaning distortion errors than good readers, and are less likely to correct their errors (Steiner, Wiener & Cromer, 1971; Weber, 1971). And poor readers tend to read in smaller units (e.g., word for word) whereas good readers read in larger units like phrases (Cromer, 1970). The reading behaviors characteristic of poor readers are problematic because they interfere with reading comprehension.

Additionally, poor readers spend significantly less time reading connected text than good readers (Allington, 1984; Gambrell, Wilson, & Gantt, 1981). As a result, poor readers are slow to develop reading skills that are facilitated by the act of reading such as developing a rich vocabulary and knowledge base, extracting

orthographic structures from print, and using context to decode unfamiliar words, define words, and correct decoding errors (Stanovich, 1986). In comparison to good readers, poor readers also spend more time on reading materials that are too difficult for them (Allington, 1977; 1983b; 1984). As a result, poor readers often experience frustration and failure when reading (Bristow, 1985; Gambrell et al., 1981), reducing further the amount of time they practice reading.

Researchers have also attempted to identify factors that contribute to the persistence of reading problems. Descriptive research on the nature of reading instruction provided in remedial and regular education classrooms indicates that there is a lack of congruency between remedial efforts and classroom reading instruction, a failure to provide additional instruction time that these students need to acquire reading skills (Allington, Stuetzel, Shake, & Lamarche, 1986; Haynes & Jenkins, 1986). It also indicates that only one-third of the remedial work is devoted to direct reading activities, and only one-fourth of this work focuses on comprehension related activities. These findings have been replicated by other researchers (e.g., Allington & McGill-Frazen, 1989; Ysseldyke, O'Sullivan, Thurlow, & Christenson, 1989). Collectively, these findings suggest that failure on the part of remedial reading programs to incorporate effective instructional reading practices may account at least partly for the prolonged reading problems experienced by low-performing readers.

## Understanding Reading Fluency

In spite of the finding that good readers tend to read larger units with fluency and accuracy (and, therefore, read with better comprehension), oral reading fluency has been largely neglected by researchers for the past several decades (Allington, 1983a). Allington has called fluency the “neglected goal” of reading research and instruction. Similarly, Anderson (1981) stated that fluency is the “missing ingredient” in reading instruction. Fluency has rarely been an instructional objective because it has long been assumed by researchers and practitioners that fluency was the immediate result of word recognition proficiency; hence researchers’ and practitioners’ focus on the development of word recognition (National Reading Panel, 2000). However, this situation is rapidly changing. Researchers have attempted to define reading fluency more appropriately. In their Dictionary of Reading, Harris and Hodges (1981) define fluency as the ability to read “smoothly, easily and readily” with “freedom from word recognition problems” and dealing with “words, and larger language units” with quickness (p.120). Moyer (1982) suggested that the two key aspects in fluency are accuracy of word recognition and rate of reading. The ultimate goal of fluency in reading is to enhance the reader’s comprehension of the text. In other words, fluent readers accurately recognize individual words and group words appropriately into meaningful grammatical units for interpretation (Schreiber, 1980, 1987). Thus, fluent readers read with relatively greater speed and accuracy, decoding and comprehending simultaneously; while dysfluent readers read word-by-word, with little understanding of the text (LaBerge & Samuels, 1974).

Researchers have also attempted to discover how reading fluency develops, and towards this end they have expressed several theories about fluency acquisition. These include resource-based theories developed by LaBerge and Samuels (1974), the interactive-compensatory model proposed by Stanovich (1980), and instance theory or information encapsulation promoted by Logan (1997).

Early understanding of what is involved in reading fluency development was largely shaped by LaBerge and Samuels' resource-based theories of reading (LaBerge & Samuels, 1974), which emphasized the importance of high-speed word recognition. In essence, LaBerge and Samuels promoted the view that skilled reading involves the reallocation of attentional capacity from lower-level word identification processes to resource-demanding comprehension functions (Fuchs, Fuchs, Hosp, & Jenkins, 2001). As a result, much conceptual thinking and empirical work has focused on word-recognition.

However, some researchers (e.g., Logan, 1985; Zbrodoff & Logan, 1986, as cited in Fuchs et al., 2001) have found that obligatory execution of word recognition develops rapidly, whereas speed and efficiency of execution continue to develop even after word recognition has become obligatory (Guttentag & Haith, 1978; West & Stanovich, 1978). This finding suggests that speed and obligatory execution can be separated from resource capacity (Stanovich, 1990). Such findings have caused a shift in recent conceptualization of reading from resource-based models toward informational encapsulation; specifically, this has caused a shift from concentration on issues of cognitive resource use to an emphasis on knowledge representation. The primary feature of information encapsulation is the development

of a knowledge base through which the reader codes information without considering alternative possibilities about what the word is or what it means (Logan, 1988, as cited in Fuchs et al., 2001; Logan, 1997). When the knowledge base is large enough and reliable enough, performance could be based entirely on memory retrieval, and the algorithm that once supported novice performance could be abandoned entirely. The key causal property is the development of a high-quality representation in memory that allows automatic access. Non-automatic performance is limited by a lack of knowledge rather than by the scarcity of resources (Logan, 1988, as cited in Fuchs et al., 2001).

Stanovich (1980), on the other hand, argued that individual differences in reading fluency development could be explained best by an interactive-compensatory model of reading performance. The interactive-compensatory model suggests that higher-level processes can actually compensate for deficiencies in lower-level processes. For example, when a word is encountered in sentence context, “bottom-up” (text-driven or word processing) and “top-down” (meaning driven or hypothesis forming) processes operate simultaneously. Thus, a reader with poor word recognition skills actually may tend to rely more on contextual factors because these factors are more accessible than text-driven factors, and provide additional sources of information. Fluent readers, by contrast, will become more dependent on context if their context-free word recognition processes are hindered by some means.

## Need to Synthesize Instructional Effectiveness

These three theories converge on the essential features of reading fluency: Fluent reading is effortless, fast, unconscious, and automatic, and acquiring it is gradual and is achieved through practice and repetition in consistent environments (Logan, 1997).

Accordingly, over the past decade, there has been a considerable amount of research in reading fluency development. Researchers have developed and tested various approaches that provide additional reading time and more intensive reading instruction, and engage students in direct reading activities to develop fluency and comprehension. These approaches include previewing, class-wide peer tutoring, computerized programs, and repeated readings, oral recitation lesson, neurological impress method, or a combined method. These interventions share similar components. All include additional practice, frequently with the same passages or word lists. Several approaches ask students to read along while they listen or to preview reading materials prior to reading the passages themselves. Additionally, many approaches ask students to try to read as fast as possible.

Many of these techniques have been shown to improve fluency and comprehension in students with and without disabilities. However, the pertinent studies have varied considerably in the persuasiveness of their respective findings: Some studies have produced inconclusive, even contradictory results; others lack methodological rigor (e.g., small or nonequivalent samples). A meta-analysis of these studies is needed to integrate evidence across these studies. A meta-analysis offers statistical analysis of a large collection of results from individual studies for the

purpose of creating generalizations (Cooper & Hedges, 1994). As Taveggia (1974) pointed out, the findings of any single study are meaningless in and of themselves. They may have occurred simply by chance. Therefore, we need to synthesize the results from individual studies to explore generalizations regarding their effectiveness. Furthermore, without such evidence, we do not know the magnitude or pattern of the effects of these interventions on fluency, or for whom they are most effective. It is also unclear which intervention approach is most or least effective in promoting fluency.

Therefore, a meta-analysis of the effects of these interventions is highly opportune. However, up to date, there have been only two pieces of research that have attempted to synthesize the effects of fluency promoting interventions meta-analytically: the National Reading Panel (NRP) report (December, 2000) and a review conducted by Kuhn and Stahl (2000). Unfortunately, these two reports each fell short of a real meta-analysis for various reasons. The NRP report did not conduct any systematic analyses on effect sizes to determine the best procedures or the features of the best procedures because there were simply not enough comparisons in either guided repeated oral reading or encouraging students to read more. The authors of the NRP report noted a lack of clear differences among fluency building procedures but nonetheless they noticed the robustness of the procedures.

Kuhn and Stahl (2000) originally planned to conduct a meta-analysis but did not for three reasons. First of all, they found few studies with control groups. Second, they found great variability in the effect sizes they calculated, ranging from .13 to 2.79. They believed that the few large effect sizes would have to be eliminated

from the analysis to avoid their having an excessive influence on the calculated effect. In the current meta-analysis, I will not eliminate very large effect sizes (outliers); instead, I will adjust them to less extreme values through a process called Winsorizing. Finally, the authors noticed that there were a number of different conditions that were used as controls, from no-treatment to having students spend an equivalent amount of time in non-repetitive reading. They believed that these different control conditions made it difficult to come up with a common metric. In my current meta-analysis, I will look at all fluency building strategies, including repeated reading and non-repeated reading strategies, and a control condition is where students received what they normally would receive (not just no treatment control).

In an earlier meta-analysis, I concluded that repeated reading was effective in improving reading fluency and comprehension in both low-achieving non-disabled readers and students with disabilities (Yang, 2000). In this second meta-analysis, as mentioned previously, I wish to examine more broadly and in greater depth the effects of various intervention approaches to increase fluency and comprehension of students with and without high-incidence disabilities at the elementary level. Specifically, instead of looking at only repeated reading, this meta-analysis examines all fluency-building interventions with all types of students at the elementary school level. The potential interventions include repeated reading, neurological impress method, oral recitation lessons, peer tutoring, and other fluency-promoting strategies. This meta-analytical review explores the relative effectiveness of many interventions, which of them are most or least effective, and for whom and under what conditions they are effective. It differs from my previous



meta-analysis in three ways. First, it includes virtually all recognized approaches that have been used to promote reading fluency at the elementary school level, rather than just repeated reading. Second, it includes studies that examined the effects of these approaches on the fluency and comprehension of all student types at the elementary level, not just remedial readers and students with disabilities. Third, it examines only experimental and quasi-experimental studies that include a control group. My prior meta-analysis, however, looked at all pre-post design studies, including one-group pre-post design studies, pre-post experimental and quasi-experimental designs studies with or without a control group. Studies with a control group are methodologically more rigorous than studies that do not involve a control group, and therefore can provide a better indicator of the effectiveness (or ineffectiveness) of these fluency building interventions. Unlike a conventional meta-analysis that only uses studies that contain one treatment and one control group, in this meta-analysis, in addition to studies that contain one treatment and one control group, I also include studies that include multiple treatment groups and a control group.

Through this effort, I wish to provide a comprehensive picture of the effects of various instructional approaches to fluency building in elementary-age students with and without disabilities. Specifically, I will examine intervention effectiveness on three reading outcomes: reading rate, reading accuracy, and/or reading comprehension. I will attempt to answer the following questions:

1. What is the average effect size representing each intervention for each reading outcome? Is the average effect size representing each intervention for each outcome significantly different from zero?
2. What is the average effect size representing each intervention for each outcome for students with disabilities? Is the average effect size of each intervention for each outcome for students with disabilities significantly different from zero?
3. What is the average effect size representing each intervention for each outcome for students without disabilities? Is the average effect size of each intervention for each outcome for students without disabilities significantly different from zero?
4. Is the variability in the distribution of effect sizes representing each intervention for each outcome heterogeneous? If the variability in the effect size distribution representing the interventions for each outcome is heterogeneous, what variables are associated with that variability? In other words, what factors are associated with or best predict the least or most effective intervention(s) to increase reading fluency?

## CHAPTER II

### LITERATURE REVIEW

In the previous chapter, the magnitude of reading problems in the United States was described. In the present chapter, I will first briefly discuss three influential theories about reading fluency development. Then I will provide a review of representative primary studies that have been conducted to improve reading fluency. Although the present meta-analysis focuses only on studies including a control group, in this chapter I will provide a comprehensive picture of fluency research at the elementary level. Specifically, I will review representative studies of all types of designs, including experimental and quasi-experimental studies with a control group, one group or multiple treatment group research without a control group, and single case efforts. By reviewing representative studies of all designs, I wish to provide a backdrop and rationale for my decision to include in my meta-analysis only studies employing a control group.

#### Theories about Reading Fluency Development

There have been several theories about the development of reading fluency or automaticity. These theories include resource-based theories of reading fluency developed by LaBerge and Samuels (1974), the interactive-compensatory model of fluency development proposed by Stanovich (1980), and instance theory or information encapsulation promoted by Logan (1997).

### *Resource-Based Theories of Reading Fluency*

Early understanding of what is involved in reading fluency development was shaped by LaBerge and Samuels' theory of automatic information processing or resource-based theories of reading (LaBerge & Samuels, 1974). This theory emphasizes the importance of high-speed word recognition, resulting in the fact that much conceptual thinking and empirical work has focused on word-recognition. According to LaBerge and Samuels (1974), the execution of a complex skill (such as reading) requires the coordination of component processes within a short time. If each component requires attention, execution of the skill will be impossible since attentional capacity will be exceeded. However, if some components can be executed automatically, there will be enough attentional capacity left for the skill to be performed successfully. This conceptualization emphasizes automaticity as an explanatory construct in developmental reading ability. The theory rests on the following assumptions. First, individuals have limited amount of attention or processing energy available for decoding and comprehending text at any given time during reading. Second, comprehension requires a great deal of attention. Third, automatic decoding does not require a lot of attention. Fourth, to develop good reading comprehension skills, a reader needs to minimize the amount of attention used for decoding and maximize the amount of attention available for comprehension.

Specifically, readers have two essential and simultaneous tasks to perform, recognizing words and deriving meaning (i.e., comprehending). When a passage contains too many words that are difficult either because they are unfamiliar or hard

to decode, readers tend to focus more attention on decoding these words. Consequently, little attention remains for comprehension. In contrast, when readers can recognize printed words automatically, word identification requires little attention and readers may focus on comprehension. In essence, LaBerge and Samuels promoted the view that skilled reading involves the reallocation of attentional capacity from lower-level word identification processes to resource-demanding comprehension functions (Fuchs, Fuchs, Hosp, & Jenkins, 2001). This and related resource-based models have been very influential and continue to influence current conceptualizations of reading fluency development (e.g., Spear-Swerling & Sternberg, 1994).

#### *Interactive-Compensatory Models of Reading Fluency*

Stanovich (1980) argued that individual differences in reading fluency development could be best explained by an interactive-compensatory model of reading performance. Interactive models (Rumelhart, 1977) assume that a word is recognized from information provided concurrently from several knowledge levels (e.g., feature extraction, orthographic knowledge, lexical knowledge, syntactic knowledge, semantic knowledge). In this model, higher level processes (e.g., semantic processes) constrain the alternatives at lower levels (e.g., orthographic or lexical knowledge) but are themselves constrained by lower-level processes. Thus, instead of being merely a data source for higher level processes, each level of processing seeks to synthesize the stimulus based on its own analysis and the limitations imposed by both higher- and lower-level analyses.

Stanovich (1980) added a compensatory component to the general interactive model to explain several findings in the literature on individual differences in the development of reading fluency. The essence of the compensatory model is the assumption that deficiencies in processes at a particular level in the processing hierarchy (e.g., orthographic, lexical, and semantic level processing) can be compensated by a greater use of information from other levels, and this compensation takes place regardless of the level of the deficiency. Therefore, the interactive-compensatory model suggests that higher-level processes can actually compensate for deficiencies in lower-level processes. For example, when a word is encountered in sentence context, “bottom-up” (text-driven or word processing) and “top-down” (meaning driven or hypothesis forming) processes operate simultaneously. Thus, a reader with poor word recognition skills actually may tend to rely more on contextual factors because these factors are more accessible than text-driven factors, and provide additional sources of information. Fluent readers, by contrast, will become more dependent on context if their context-free word recognition processes are hindered by some means. Support for the interactive-compensatory model comes from studies showing a pattern of greater contextual reliance by less fluent readers (e.g., Leu, DeGross, & Simons, 1986; Stanovich & Stanovich, 1995; Stanovich & West, 1981).

#### *Instance Theory / Information Encapsulation of Reading Fluency*

Within resource-based models of reading, researchers tend to equate different dimensions of automatization, such as obligatory execution (i.e., the

tendency for an automatized process to execute regardless of where conscious attention is directed), speed, and capacity-free processing (Fuchs et al., 2001). Some researchers (e.g., Logan, 1985; Zbrodoff & Logan, 1986, as cited in Fuchs et al., 2001), however, have shown that obligatory execution of word recognition develops rapidly, but that speed and efficiency of execution continue to develop even after word recognition has become obligatory (Guttentag & Haith, 1978; West & Stanovich, 1978). This finding suggests that speed and obligatory execution can be separated from resource capacity (Stanovich, 1990).

Such findings have caused a shift in recent conceptualization of reading from resource-based models toward informational encapsulation, another characteristic associated with the automaticity concept (Logan, 1997). Specifically, this has caused a shift from concentration on issues of cognitive resource use to an emphasis on knowledge representation. The primary feature of information encapsulation is the development of a knowledge base, through which the reader codes information without considering alternative possibilities about what the word is or what it means (Logan, 1988, as cited in Fuchs et al., 2001; Logan, 1997). According to Logan, performance is automatic when it is based on retrieval of past instances – memories of past solutions to task relevant problems, rather than algorithmic computations (such as by thinking or reasoning). He argues that the greater the number of task-relevant instances in memory, the more likely automatic performance. When the knowledge base is large enough and reliable enough, performance could be based entirely on memory retrieval and the algorithm that once supported novice performance could be abandoned entirely. The key causal property is the

development of a high-quality representation in memory that allows automatic access. Non-automatic performance is limited by a lack of knowledge rather than by scarcity of resources (Logan, 1988, as cited in Fuchs et al., 2001).

These three theories converge on the essential features of reading fluency: Fluent reading is effortless, fast, unconscious, and automatic. Acquiring it is gradual and is achieved through practice and repetition in consistent environments (Logan, 1997). Empirical research on reading fluency acquisition both borrows support from and provides support for these theories.

#### A Review of Primary Studies Conducted to Improve Reading Fluency

There has been a considerable amount of research on effective techniques for improving reading fluency. Studies have found support for the effectiveness of some interventions, such as previewing, tutoring, class-wide peer tutoring, cross-age tutoring, computerized programs, repeated reading (Mastropieri, Leinart, & Scruggs, 1999), neurological impress method (Bos, 1982), and oral recitation lesson (ORL) (Aslett, 1990; Hoffman, 1987; Hoffman, Isaacs, Roser, & Farest, 1989; Nelson & Morris, 1986; Reutzel & Hollingsworth, 1993). Research has also examined the effects of rapid decoding training on reading fluency development (Fleisher, Jenkins, & Pany, 1979; Grant & Standing, 1989; Gun, Biglan, Smolkowski, & Ary, 2000; Marston, Deno, Kim, Diment, & Rogers, 1995). In the following section, I review studies that have used each of these approaches.



## *Repeated Reading*

### *Single Case/Subject Studies*

Subjects in these studies were either at-risk children or students with disabilities. There are two variations in the method of repeated reading, namely, repeated reading itself (e.g., Swain & Allinder, 1996; Turpie & Paratore, 1995; Weinstein & Cooke, 1992) and a modified version of repeated reading called reading to read (RTR) (e.g., Bolton, 1991; Boyer, 1991; 1992; Cottingham, 1996; Morris, 1995; Prestridge, 1996; Tingstrom, Edwards, & Olmi, 1995).

Swain and Allinder (1996) used a multiple baseline across subjects to examine the effects of repeated reading on the reading performance of three second-grade students with learning disabilities. Two types of curriculum-based measurement (CBM) procedures, Oral Reading and Computer Maze, were used to measure student performance. Results demonstrated that only one student showed a dramatic level change between conditions and only in oral reading CBM. In terms of slope, two students showed improvements during the intervention in the oral reading CBM. However, slopes for all three students declined during intervention for the maze CBM assessment. In terms of weekly gains, the mean average weekly gain for oral reading CBM for baseline was  $-.46$ . When the intervention of repeated reading was introduced, the mean average weekly gain was 2.33 words, far above the goal line of 1.5 words per week. However, the average weekly gain for maze CBM was less than zero. It appeared that oral reading CBM was positively affected by the intervention of repeated reading, whereas maze CBM was not. The authors

maintained that repeated reading increased students' reading rates, which in turn affected oral reading CBM, but these improvements did not transfer to maze CBM.

Weinstein and Cooke (1992), on the other hand, used a multitreatment, single subject design (ABACA) to compare the effects on fluency of two types of mastery criteria for repeated reading. The two interventions were (a) repeated reading with a mastery criterion of 90 correct words per minute (or the fixed-rate criterion phase), and (b) repeated reading with a mastery criterion of three consecutive fluency improvements (or the improvement criterion phase). Four students with LD, aged 7 to 10, participated in the study. All students made gains in their mean correct words per minute (CWPM) with the repeated reading technique, regardless of the type of criterion. Comparison of intervention phase means with preceding baseline means showed gains ranging from 16.1 to 39.4 correct words per minute. The mean CWPM for all students in the fixed rate criterion phase was 64.9, with a range from 62.3 to 68.5. the mean CWPM for the successive improvements criterion phase was 54.5, with a range from 50.4 to 69.0. The relationship between the mean gains and the previous baseline is reflected in the percentage of gain which averaged 62% (mean gain of 24.8 WCPM) for the fixed rate phase and 58% (mean gain of 20.0 CWPM) for the improvements phase.

The efficiency of each phase was determined by comparing the number of rereadings required to reach criterion, which favored the improvements phase; twice as many rereadings were required to reach criterion in the fixed-rate phase as in the improves phase. The extent to which fluency gains transferred or generalized to unpracticed passages was determined by comparing means from Baseline 1 to

Baseline 2 and from Baseline 2 to Baseline 3. The improvement criterion phase also showed advantage over the fixed-rate criterion with respect to generalization of fluency.

In reading to read (RTR), several procedural changes were incorporated with the intent of improving the effectiveness of traditional repeated reading methods. Specifically, changes included immediate corrective feedback, performance feedback, verbal reinforcement, student (self) charting progress, and progression through the curriculum based upon acquisition of prerequisite skills and skill mastery. Bolton (1991) employed a multiple-baseline design across subjects to examine the overall effectiveness of RTR on the fluency of poor readers. Six third-grade students participated in the study. Results were indicative of improved oral reading fluency and suggested that gains in fluency differed as a function of entry level achievement. RTR was shown to be more effective than the basic repeated reading procedure with regard to improved fluency.

Boyer (1991), using a multiple baseline across subjects, also investigated the effectiveness of RTR relative to improved oral reading fluency. Four second-grade students from the lowest reading group in a classroom participated. Each student met individually with the researcher for approximately 34 sessions. Results showed that oral reading performance improved significantly for all participants. Boyer (1992) next examined the effects of RTR on passage comprehension. Generalization of fluency gains achieved during RTR intervention to unpracticed passages from a parallel reading series as well as to a norm-referenced standardized achievement test was also assessed. Eight poor readers in first and second grade participated in

the study. A modified multiple baseline design was used to compare the performance of these students with a traditional repeated reading procedure and with RTR. All subjects demonstrated gains in oral reading fluency and passage comprehension with both traditional repeated reading and RTR; however, RTR gains were significantly superior to those made during traditional repeated reading intervention. RTR was also more effective relative to generalization of oral reading fluency gains to unpracticed passages as well. These results were replicated in Prestridge's (1996) and Cottingham's (1996) studies.

To summarize, these single case/subject studies demonstrated that repeated reading (and its variant RTR) was effective in improving the oral reading fluency of at-risk children and students with disabilities. Considering the fact that the number of subjects available in special education is limited, single-case/subject designs can be a meaningful approach to demonstrating the effectiveness of repeated reading as a viable remedial instruction method in reading. However, due to the limitations inherent in this type of design, i.e., low external validity, it is impossible to generalize the findings to other student populations and settings.

#### *Group Design Studies Without A Control Group*

These studies used either a one-group pretest posttest design or multiple treatment group design. The one-group pretest posttest design studies examined the effects of repeated reading on fluency and/or comprehension without including a control group (e.g., Goldstein, 1999; Herman, 1985; Levy, Nicholls, & Kohen, 1993). The multiple treatment group design studies compared the effects of different variations of repeated reading or other interventions without a control group (e.g,

Dowhower, 1987; Homan, Klesius, & Hite, 1993; Komisar, 1999; O'Shea, Sindelar, & O'Shea, 1985, 1987; Rasinski, 1990; Selvey, 1990; Sindelar, Monda, & O'Shea, 1990; Stoddard, Sindelar, O'Shea, & Algozzine, 1993; Vaughn, Chard, Bryant, Coleman, Tyler, Linan-Thompson, & Kouzekanani, 2000).

*One group pretest posttest studies.* Herman (1985) attempted to determine whether reading gains were limited to practiced passages or whether they also transferred to unfamiliar passages. Subjects were 80 elementary-aged students from remedial reading and math programs, with an oral reading rate of 35 to 50 words per minute. They repeatedly read five passages of their own choosing until a criterion reading speed was reached on each. The initial and final reading scores on two of the five passages were analyzed. Results indicated that students' reading rate, accuracy, and comprehension improved on both practiced and unpracticed passages. But since there was no control group, it was difficult to determine whether the findings were due to repeated reading or other variables. Further, the small sample size limited the generalizability of the findings to other students.

Goldstein (1999) also used a one group pretest posttest design to examine the effects of repeated reading on the oral reading fluency of 11 first and second grade poor readers. Students were engaged for 4 months in repeated reading of text, auditory modeling, supported reading techniques, and direct instruction. A comparison of pretest and posttest results showed that the students improved words read correctly per minute and fluency reading skills according to a fluency rating system. However, as in the Herman (1985) study, the lack of a control group and the

small sample size made it difficult to determine the specific effects of the intervention.

*Multiple group no control studies.* Sindelar, Monda, and O'Shea (1990) investigated the importance of repeated reading on the reading fluency of elementary-aged students with LD and without disabilities. All students participated in non-repeated reading and repeated reading conditions. During the non-repeated reading condition, students read one of two experimental passages once. During the repeated reading condition, they read the second passage three times. Sindelar et al. reported that reading rate increased significantly from one to three readings, an occurrence that brought instructional-level readers to near mastery-level performance. The authors concluded that the effects of repeated readings were comparable for students with and without LD. However, because the authors did not include a control group, they could not rule out alternative explanations (e.g., natural maturation) for improved fluency and comprehension. Further, it was unclear whether the two experimental conditions occurred on the same day or different days. If they occurred on the same day, then repeated reading effects might have been due to four rather than three readings. If the experiment occurred across days, it would have been helpful to clarify the amount of time that elapsed between pretest and posttest.

In a similar study, Stoddard et al. (1993) reported that the reading rate of fourth and fifth grade students improved significantly from one reading to three readings to seven readings. Student comprehension improved significantly when the

number of readings was increased from one to three readings, but not from three to seven readings.

O'Shea, Sindelar, and O'Shea (1985) also investigated the effects of reading a passage one, three, and seven times, and of repeated reading with different attentional cues on the reading fluency and comprehension of 30 third-grade general education students. Students were randomly assigned to one of two experimental conditions, cue for fluency or cue for comprehension. In the cue for fluency condition, students were encouraged to read as quickly and accurately as they could, whereas in the cue for comprehension condition, students were encouraged to remember as much of the story as they could. All students read one passage once, a second passage three times, and a third passage seven times. Reading comprehension (the proportion of propositions recalled) and fluency (the rate at which passages were read) were assessed on the final reading of each passage. Effort was made to ensure the passages were of comparable difficulty and the order of passages and repeated reading levels were counterbalanced across subjects.

Results suggested that improved fluency and comprehension occurred with each repeated reading for all subjects. Students cued to read for meaning made greater comprehension gains than students cued to read for fluency. Similarly, students cued to read fluently made greater fluency gains than students cued to read for meaning. However, study findings are limited to students with average and good reading skills. Further, although the authors contended that differences between groups were a function of the type of verbal cue, their conclusion may be questioned because additional variables were changed between groups. Students in

the fluency group were made aware that their oral reading was being timed, and were led to believe that their oral reading was being tape-recorded, and that their oral retell was not being tape-recorded. Conversely, students in the comprehension group (a) were not made aware that their oral reading was being timed when in fact it was timed with a concealed wrist watch; (b) were not told that their oral reading was being taped-recorded; and (c) were told that their oral retell was being tape-recorded. As a result, the factors that accounted for the differences between the two groups cannot be determined. And, of course, the study did not include a control group. Without the benchmark provided by a control group, we cannot know for sure whether the two conditions were influential above and beyond the students' regular program or natural progress.

In another study, O'Shea, Sindelar, and O'Shea (1987) examined the effects of repeated reading and attentional cues on 32 students with LD. The two conditions were the same as in the O'Shea et al. (1985) study. Results showed a statistically significant main effect for attentional cues on reading comprehension, but not on reading rate. Students who were cued to read for comprehension retold more of the stories but read no slower than students cued to read quickly and accurately. The authors concluded that students responded to the comprehension cue, but not to the fluency cue. In addition, a statistically significant main effect was found for repeated readings on reading rate and comprehension. Students read more fluently on seventh than third readings and on third than first readings. For comprehension, students retold more of the stories after three readings than after one, but no significant difference was found between the third and seventh readings. But again,



the lack of a control group, or a repeated reading group without attentional cues, limited the interpretation of the results. Without a control group (that did not receive a repeated reading treatment), the authors may not determine with certainty that repeated reading yielded significant results; and without a-repeated-reading-only group, the authors may not conclude that the differences were due to attentional cues. Repeated reading might have produced the improvement in the students' performance.

Rasinski (1990) investigated the relative effects of repeated reading and repeated listening while reading (RLWR) on reading fluency. Rasinski conceptualized repeated reading somewhat differently from previously discussed investigators in that students read the same passage once daily over a period of days rather than reading a different passage several times daily over a period of days. Twenty 3<sup>rd</sup>-grade students were paired by reading level (high, medium, and low) and assigned to one of two groups. Students assigned to Group 1 participated in repeated reading during the first treatment phase and did RLWR during the second treatment phase. Students in Group 2 participated in the same treatments but in reverse order. Two passages of comparable difficulty were used; one for each treatment condition. Each treatment phase lasted 4 days. On Day 1, students' oral reading fluency was assessed. On Days 2 and 3, they participated in the treatments using the same passage from the previous day. Students in the repeated reading condition read the passage aloud to their teacher. Students in RLWR listened to their teacher read the passage aloud. Order of passage presentation and treatment was counterbalanced. Results indicated that both methods were equally effective in

improving reading fluency. However, subjects were not randomly assigned to treatment groups, and there was no control group.

Researchers investigated the effects of other variations of repeated readings as well. For example, Dowhower (1987) explored the effects of independent (unassisted) versus read-along (assisted) repeated reading with poor readers in second grade. Seventeen students were randomly assigned to one of two treatment groups (but not control group). In the assisted repeated reading group, subjects rehearsed with a tape recorded passage, while in the unassisted repeated reading group, subjects rehearsed independently. Both conditions required students to practice until they met a criterion of 100 words read per minute. Four findings were reported. First, increased reading rate led to statistically significant increases in reading accuracy and comprehension on practiced passages. Second, reading rate improved significantly from the first reading of the practiced passage to the reading of the unpracticed passage. Third, significant incremental gains in rate and accuracy across the five initial readings of the practice passages were obtained. Fourth, the number of rereadings necessary to reach criterion decreased incrementally across five passages.

In summary, among this group of studies, variations of repeated reading were compared to each other or to a non-repeated reading intervention. These studies purportedly showed that repeated reading and its many variations were effective in improving fluency and/or comprehension, but they all suffered from a common problem, i.e., the lack of a control group. Another typical limitation was that researchers used small samples, limiting the generalizability of their findings.

### *Controlled Studies*

These studies are grouped in two categories. The first group consists of studies that used a repeated measures design (e.g., Bohlen, 1988; Rashotte & Torgesen, 1985). The second group of studies, all included in the present meta-analysis, compared repeated reading with a no-treatment control group (e.g., Conte & Humphreys, 1989; Hannah, 1994; Koch, 1984; Koskinen & Blum, 1984; Laffey, Kelly, & Perry, 1979), or contrasted repeated reading with a different treatment and a control group (e.g., Cohen, 1989; Collins, 1994).

*Repeated measures design studies.* Rashotte and Torgesen (1985) hypothesized that the amount of word overlap between passages used during repeated reading could explain the transfer effect of reading gains from practiced to unpracticed passages. They used a repeated measures design and investigated the overall and relative effects of repeated reading with word overlap, repeated reading without word overlap, and non-repeated reading methods on the reading fluency and comprehension performance of 12 students with learning disabilities. Students participated in each experimental condition for seven days. Treatment order was counterbalanced between subjects. In the two repeated reading conditions, students read one new passage four times daily, whereas in the non-repeated reading condition, students read four different passages once daily. Fluency was measured by the accuracy (number of errors per passage) and speed (number of words read per minute) with which students read. Comprehension was measured by the number of literal comprehension questions answered correctly.

Results indicate that students in the repeated reading with word overlap group experienced greater improvements in reading speed than students in the repeated reading without word overlap. Further, when stories shared few words, repeated reading was not more effective than an equivalent amount of non-repeated reading. Hence, improvement in reading speed was dependent on the amount of shared words among stories. In contrast to other repeated reading studies, improvement in reading accuracy and comprehension was not found. This may have been due to ceiling effects because most students had high pretest comprehension scores and were highly accurate on pretest.

Bohlen (1988) examined whether repeated reading of stories would affect fluency among five third-grade students in the “low” reading group in a self-contained classroom who read 10 stories during the study. Instruction during the first five stories followed the traditional format, whereas instruction during the second five stories included repeated reading with taped recordings of the stories. Pretest and posttest scores were obtained for both groups. Bohlen conducted two separate t-tests on the pretest and posttest scores, one for the repeated reading group and one for the control group, and found no statistically significant difference for either group. The author concluded that the repeated reading intervention did not improve the fluency of these slow readers. However, the data analysis procedures were questionable as the author did not seem to consider sample size and design. A repeated-measures ANOVA with a capacity to explore the interaction between intervention and time of testing would have represented a more powerful analysis. Alternatively, Bohlen could have used a paired t-test on the gain scores between

pretest and posttest scores for the two groups. Since the author provided pretest and posttest raw scores for both groups, I conducted a repeated measures ANOVA on the pretest and posttest scores and a paired t-test on the gain scores. Both tests were statistically significant, favoring repeated reading intervention. The study failed to yield sufficient evidence to conclude that repeated reading did not help students improve their fluency.

*Controlled group design studies.* Koskinen and Blum (1984) investigated the effectiveness of repeated reading as a classroom strategy for enhancing the fluency of 32 below-average third grade readers enrolled in six public schools. Teachers and their intact group of students were randomly assigned to one of the two treatment groups: repeated reading or study activities. Statistically significant differences were found favoring the repeated reading program on oral reading fluency. In addition, subjects in the repeated reading condition made significantly fewer semantically inappropriate miscues.

Conte and Humphreys (1989) conducted a study in which 13 9 to 13 year old poor readers listened to and read audiotaped stories until the passages could be read fluently without the tape. A same-age control group ( $n = 13$ ) with similar reading difficulties received an alternative reading program equal to what was given to the experimental group (with respect to the content or curriculum), but different in terms of passage reading exercises. The experimental group did repeated reading using taped material for 15-20 minutes while the control group read books from basal readers. This procedure was implemented three times a week for 10 weeks. Results showed a positive effect of repeated reading on oral reading (as measured by grade

equivalence scores) and on isolated word reading as measured by the Boder test (Boder & Jarrico, 1982). The experimental children also showed larger although not statistically significant gains than control children on oral and silent reading on Ekwall Reading Inventory (Ekwall, 1979), Woodcock Passage Comprehension (Woodcock, 1973), and Boder Reading Grade measures. The control group made greater gains only on Word Attack subtest of the Woodcock Reading Mastery Tests (Woodcock, 1973). However, unlike Koskinen and Blum's (1984) study, this study found no positive effect of repeated reading on reading speed. In fact, the control children read significantly faster than the experimental children. The authors suggested that the content of the two reading programs may have led to this result. That is, the tape recordings used in the repeated reading condition may have given the children a stronger orientation toward passage meaning while reading aloud. However, the absence of tape-recorded materials in the control condition may have inadvertently fostered the development of word analysis skills in students who were less capable when faced with passages of text.

In a similar fashion, Laffey, Kelly, and Perry (1979) examined the effects of repeated reading of taped literature on the reading fluency of culturally different remedial readers in grades 5 and 6. Twenty students were selected using a matched pairs design. The experimental treatment for 10 students consisted of reading while listening to the tapes and performing various comprehension activities. During 13 weeks of treatment, the experimental students spent approximately 15 minutes per day listening to the stories. After two or three times, they were encouraged to read along orally as they listened. This was repeated until they read the story fluently.

Comprehension, vocabulary, and accuracy scores were consistently higher for the experimental children than for the controls. However, the only statistically significant difference was on the comprehension scores. As was the case with related studies, the sample size in Laffey et al.'s study was small and caution is needed when interpreting results.

Koch (1984), however, used a simple one-way posttest-only control design and examined the effects of different numbers of repetitions on the reading accuracy, rate and comprehension of second grade average and above-average readers. Forty-eight students were randomly assigned to one of four conditions: two repetitions, four repetitions, six repetitions, and control. Students were tested three times, immediately after intervention (first immediate test), the day after the first immediate test (second immediate test), and delayed test (7 to 9 days later). In terms of words read correctly, the three experimental groups read more words correctly than the control, but there was no statistically significant difference among the three experimental groups. In terms of rate (words per minute), the two repetitions group read faster than the other groups on all three tests, differing significantly from the six repetitions and the control groups. The six repetitions group was the most accurate but also the slowest readers. The fastest readers (the two repetitions group) had either the second or third best word recognition accuracy. There were no differences among groups on their ability to answer either text explicit or text implicit questions. Therefore, statistically significant differences among the groups for word recognition accuracy and rate were not automatically accompanied by significant changes in subjects' explicit or implicit comprehension. Koch

suggested the difficulty of the test passages may have explained the lack of effects on comprehension. In general, study findings were mixed, with the only clear result that the six-repetition group had the best word recognition accuracy, the slowest rate, and the poorest text comprehension.

To summarize, even among this group of relatively well controlled studies, results were mixed and inconclusive due to variations in how the researchers implemented repeated reading intervention. It would therefore be meaningful to examine the relationships between these variations and the different effects associated with them, which was one of the purposes of the present meta-analysis.

#### *Classwide Peer Tutoring/Reading PALS*

Peer tutoring has been recommended as a successful way of providing additional practice for students (Scruggs & Mastropieri, 1998). With a peer tutoring approach, one half of the class can be reading at a particular time, while the other half is actively monitoring their performance, in contrast to whole class (or even small group) reading when fewer students can be actively reading simultaneously.

Two successful peer-mediated interventions are Peer-Assisted Learning Strategies (PALS) (e.g., Allor, Fuchs, & Mathes, 2001; Fuchs et al., 2001; Fuchs & Fuchs, 1998; Fuchs, Fuchs, & Burish, 2000; Fuchs, Fuchs, Mathes, & Simmons, 1997; Mathes & Babyak, 2001; Mathes, Howard, Allen, & Fuchs, 1998; Simmons, Fuchs, Fuchs, Mathes, & Hodge, 1994, 1995) and classwide peer tutoring (CWPT) (Greenwood, Delquadri, & Carta, 1988). Mathes and Fuchs (1993) investigated the effects of two conditions of CWPT (repeated reading vs. sustained reading) with a



control condition on the reading skills of students with LD in grades 4 to 6. Mathes and Fuchs reported that sustained reading within a CWPT procedure positively influenced reading fluency more than typical reading instruction. Results also indicated the repeated reading condition was not superior to the sustained reading condition or typical reading instruction in promoting fluency. In a related study, Simmons, Fuchs, Fuchs, Hodge, and Mathes (1994) examined the effects of four variations of CWPT in classrooms that included students with LD, low-achieving students, and average-achieving students. Students in all CWPT conditions made more improvements in fluency than students in a traditional instruction condition. However, few differences with respect to reading fluency were noted among the experimental conditions.

PALS, which was initially developed for students in grades 2-6 (Fuchs, Fuchs, & Mathes, 1997), has been extended to kindergarten and first grade (Fuchs, Fuchs, Thompson et al, 2001; Fuchs, Fuchs, et al. 2001) classrooms. PALS was also extended to more advanced reading instruction in high school classrooms (Fuchs, Fuchs, Thompson et al., 2001). Results of the studies at kindergarten and first grade indicate that PALS is a viable approach to improving reading fluency in early reading.

Findings generally indicate that peer tutoring improves reading fluency among many poor readers with or without disabilities presumably because it provides students with many more opportunities to read aloud as well as to read in contexts that encourage reading quickly. The studies in this category are generally well

conducted, with sound research design, relatively long treatment durations, and reliable dependent measures.

### *Previewing*

Previewing involves pre-exposure to reading passages before they are formally read. Students can preview the material aloud, silently, or by listening to the teacher or a higher-performing peer read the material (Rose, 1984a; Sindelar, 1987). Previewing is similar to repeated reading, but in some variations, such as listening to a teacher or peer reading, students can gain exposure to vocabulary, phrasing, and emphasis before reading the text themselves. Moreover, previewing may make it easier to anticipate and predict more difficult words. Research completed in this area was conducted by Rose and colleagues in a series of studies that examined the overall and relative effects of silent previewing and listening previewing on the oral reading fluency of school aged children (Rose, 1984a, 1984b, 1984c; Rose & Beattie, 1986; Rose & Sherry, 1984).

Studies were conducted with various student samples and procedures. Previewing effects were investigated with adolescent students with learning disabilities (Rose & Sherry, 1984), and elementary-age students with “mild” learning disabilities (Rose, 1984a), “severe” learning disabilities (Rose, 1984b), and behavior disorders (Rose, 1984c). Rose (1984a) compared the effects of silent previewing and previewing with listening with six students with LD using an alternating treatment design. Rose found that both previewing procedures increased reading fluency relative to a baseline (no previewing) condition, and that the highest levels of

performance were achieved in the previewing with listening condition. Rose and Beattie (1986) compared the effects of teacher-directed and taped previewing on students with LD. Students in both conditions outperformed controls, and the listening previewing condition was again associated with the highest oral reading rates. Salen and Nowak (1988) investigated the effects of a peer previewing procedure on reading rate and accuracy for three elementary school students with LD. The authors concluded that the procedure resulted in a decrease in errors and an increase in reading fluency. Findings from these studies suggest that various types of previewing activities improve reading fluency.

#### *Neurological Impress Method*

Neurological impress method (NIM) is a system of unison reading whereby the student and teacher simultaneously read aloud at a rapid rate (Heckelman, 1969). The teacher uses a finger as a locator and reads into the ear of the student. The goal is to read as many pages of material as possible in the time available and without causing physical discomfort. The teacher does not attempt to teach sounds or word recognition. The few studies that have investigated the efficacy of the method have produced inconclusive findings. Whereas several teams of researcher have reported that the NIM improved reading rate and accuracy (Langford, Slade, & Barnett, 1974; Hollingsworth, 1978), and reading comprehension (Embry, 1968; Heckelman, 1969), additional research has shown that the approach failed to produce gains in reading fluency (Hollingsworth, 1970), word recognition, and reading comprehension (Lorenz & Vockell, 1979).

These contradictory research findings leave practitioners without a clear empirical basis to evaluate the applicability of the NIM for classroom instruction. Underscoring this point is that the research on NIM is typified by a lack of control groups and small sample sizes. Other researchers (e.g., Bos, 1982) recommended using a combination of the repeated readings methods and the NIM, referred to as the repeated choral reading. This method also combines aspects of oral previewing and repeated reading by incorporating a model, repetition, and mastery criteria (Mathes, Simmons, & Davis, 1992). It may have promise, but has not been adequately researched.

#### *Oral Recitation Lesson (ORL)*

In the ORL, teachers first model the reading of a text, then students practice aloud assigned parts together and independently in preparation for a scheduled recitation. ORL has been recommended by researchers and others as an important means of providing fluency instruction in the regular classroom reading curriculum (Rasinski, 1990; Reutell & Hollingsworth, 1993)). The use of such a lesson structure may provide a direct test of the effect of fluency development on reading comprehension.

Hoffman (1987) and Nelson and Morris (1986) found ORL useful for helping at-risk readers make substantial gains in reading achievement. Hoffman, Isaacs, Roser, and Farest (1989) noticed that first and second-grade students taught with ORL made substantial (though not significant) gains in reading fluency. In another study, Aslett (1990) discovered that fourth-grade students participating in fluency

training that included ORL improved their performance on measures of comprehension that included the ability to map the text and write summaries, but not on a comprehension test using true-false items. Reutzell and Hollingsworth (1993) found that the ORL was an effective means of developing second-grade students' oral reading fluency as well as comprehension.

### *Computer-Assisted Instruction*

Another possible method for increasing practice to promote reading fluency is through computer-assisted instruction (Carver & Hoffman, 1981). Jones, Torgesen, and Sexton (1987) used the Hint and Hunt program to explore the effects of computer-guided practice on reading fluency and concluded that the experimental groups considerably improved reading fluency and accuracy and generalized to reading similar words in context.

Other researchers (e.g., Reitsma, 1988) compared the effects of computer-based speech feedback with guided reading and reading-while-listening conditions. They reported significant increases in reading fluency between treatment condition and the baseline but equivocal results among treatment conditions. Royer and Sinatra (1994) and Royer (1997) described a computerized reading assessment linked to fluency building interventions emphasizing reading sight words from word lists quickly and accurately. Preliminary results indicated that students' reading performance, including word reading rate, improved substantially as measured on the computerized reading tests. This suggests that practice in reading sight word lists can positively affect oral reading fluency for students with reading difficulties.

### *Decoding Training*

There have been two views regarding the relationship between decoding (the ability to identify a written word) and comprehension (the ability to extract meaning from the printed word). One view holds that comprehension is not a direct function of the ability to organize words in meaningful sequences (Smith & Holmes, 1970-1971; Oaken, Wiener, & Cromer, 1978). The opposite view suggests that decoding is the primary component of reading comprehension (Perfetti & Hogaboam, 1975; Isakson & Miller, 1976). It is also argued (LaBerge & Samuels, 1974) that a fluent reader not only possesses skills which include decoding, but has mastered them to the point of automaticity so that they require little conscious attention. Perfetti and Hogaboam (1975) may possess decoding skills that are accurate but slow and not yet automatic.

Studies to determine whether training poor readers in rapid decoding can improve reading fluency and comprehension have produced discrepant results. Blanchard (1981) reports unpublished studies suggesting that very poor readers do benefit from rapid decoding training, unlike moderately poor readers. However, Fleisher, Jenkins, and Pany (1979) found no benefit from rapid decoding training (using either isolated words or phrases) upon comprehension, though the training did successfully increase decoding rate. There are several methodological limitations in the study. Notably, the training in rapid recognition of word phrases produced no effect on contextual reading rate. Since reading was tested in context, increased decoding speed in context appears to be an essential prerequisite for a valid examination of the relationship between speeded word recognition and

comprehension. Another problem is that the average reading level of the poor readers trained in rapid recognition of individual words was 2.8 while the two test passages were 6.3 and 7.1 reading levels. It is possible that the vocabulary of the test passages was not appropriate to the subjects' language experience.

Grant and Landing (1989) extended the work of Fleisher et al. to provide a more thorough test of the effect of training in rapid decoding on reading performance. Eighty subjects aged 12 to 15 were identified as good or poor readers on the basis of reading rate and comprehension scores on the Stanford Diagnostic Reading Test (SDRT). Good readers were those at or above the 7<sup>th</sup> stanine in both skills, while poor readers scored at or below the 4<sup>th</sup> stanine. Two conditions of twenty students each were trained in rapid decoding of words presented either individually or in context. Twenty good readers and twenty poor readers provided additional untrained controls groups. Results indicate that both conditions significantly increased their reading rates. Improved comprehension was noted for all four groups due to a general practice effect. However, context-trained students did not differ from the subjects trained with individual words on either reading rate or comprehension.

### *Specific Instructional Methods*

In addition to the above-mentioned specific fluency building strategies, researchers have examined the effects on reading of strategies such as explicit teaching, effective teaching, direct instruction, and reciprocal teaching. For example, Simmons, Fuchs, Fuchs, Mathes, Hodge (1995) examined the effects of explicit

teaching and peer tutoring on the reading achievement of students with learning disabilities (LD) and non-disabled, low-performing readers in academically integrated general education classrooms. Teachers were assigned randomly to explicit teaching, explicit teaching plus peer tutoring, or control condition. Teachers were trained to organize instruction according to teacher presentation, guided practice, and independent practice phases. Teachers were encouraged to consider the overall time available for instruction, review prerequisite skills, frame the lesson by telling students what will be studied and why it is important, model the target skill, conduct guided practice with multiple examples, provide immediate and substantive corrective feedback, prepare students for independent practice, actively monitor students during independent practice, and review critical information at the close of the lesson. Treatment lasted sixteen weeks. Results indicate that students in the explicit teaching plus peer tutoring condition scored significantly higher on reading fluency and comprehension measures than did explicit teaching or control students.

Marston, Deno, Kim, Diment, and Rogers (1995) examined the effects of direct instruction, effective teaching, computerized decoding training, peer tutoring, reciprocal teaching on reading performance of students with mild disabilities. Results show that direct instruction and computerized decoding training had statistically higher scores than the control, and among the treatment groups, the computerized decoding training group did significantly better than the peer tutoring and effective teaching groups. Direct instruction students did significantly better than those in peer tutoring group. Further, reciprocal teaching group and direct instruction group did significantly better than the peer tutoring group.



Some researchers also looked at the effects of some commercially produced reading programs, such as Accelerated Reader and Reading Recovery, on the reading fluency of students. For example, Denton (1997) examined the effects of Reading Recovery on the reading performance of at-risk first-grade students. Student performance was measured using an informal survey of oral reading fluency, accuracy, and comprehension. The author concluded that while the program helped students to achieve an accelerated rate of growth in their literacy learning and to read accurately and with average comprehension, the treatment group students did not achieve expected normative benchmarks in oral reading fluency on the norm-referenced measure, and they did not seem to attain the average level of performance of their peers on literacy measures.

Vollands, Topping, and Evans (1999) conducted a quasi-experimental action research on the effects of the Accelerated Reader. They looked at the formative effects on reading achievement and motivation in two schools in severely socio-economically disadvantaged areas. The results suggest that the program, even when less than fully implemented, yielded gains in reading achievement for these at-risk readers that were superior to gains from regular classroom teaching and an alternative intensive methods, even with less time devoted to class silent reading practice than in comparison classes.

### *Summary of the Interventions*

Rasinski (1989) has identified six principles that can guide the development of appropriate fluency instruction in the classroom. These principles include modeling

for students, direct instruction and feedback, providing support for students while they are reading (e.g., choral reading and reading while listening), repeated readings of one text, cueing phrase boundaries in texts, and providing students with easy materials.

The fluency-building approaches reviewed share some similar components: All include additional practice, frequently with the same passages or word lists. Several approaches ask students to read along while they listen or to preview reading materials prior to reading the passages themselves. Many approaches ask students to read as fast as possible. A majority of the approaches have been shown to improve reading fluency and comprehension in elementary students with and without disabilities. However, the pertinent studies have varied considerably in the persuasiveness of their respective findings. Some studies have produced inconclusive, even contradictory results; many studies lack methodological rigor.

#### *Treatment Effects Do Not Transfer*

In the case of repeated reading, two major approaches have been advocated. In one, where non-taped materials are used, the student repeatedly reads a passage until a fluency criterion is achieved (Samuels, 1979). In the other, in which taped material is used, the student first listens to an audiotape recording of the passage, then reads along with the tape, and then attempts to read unassisted until a fluency criterion is reached (Chomsky, 1978). Both approaches have been shown to be an effective component of a remedial reading program (e.g., Koskinen & Blum, 1984; Dahl, 1979; Conte & Humphrey, 1984).

Nonetheless, most studies indicate that the transfer of gains to novel material

is limited. For example, Rashotte and Torgesen (1985) found statistically significant effects of non-taped repeated reading on reading speed only on passages that shared similar content with the training passages. Carver and Hoffman (1981) reported significant transfer to passages that were presented in a similar format (multiple choice, cloze passages) to the training passages. Herman (1985) reported gains in reading speed, word recognition, and a decrease in miscues on novel passages. However, since Carver and Hoffman (1981) and Herman (1985) used no control group, the gains may not necessarily be attributed to non-taped repeated reading. Carbo (1979) found a gain of three months in word recognition after implementing non-taped repeated reading with eight students over a three-month period, but, again, no controls were used.

ORL and paired repeated reading (Kokinen & Blum, 1986, 1987) point to the potential of fluency instruction for strengthening reading development. The studies, however, involved treatments of brief duration, did not use a broad range of fluency principles, and did not report results in a quantifiable format. Hollingsworth (1970) used a version of the neurological impress method (NIM) that was very similar to taped repeated reading. Subjects who were at or near grade level listened to and read along with tape recordings of textual material in a group setting. Subjects in experimental and control were matched on both IQ and reading measures. Results indicated no effect of the experimental treatment on Gates-McGinitie vocabulary, comprehension, speed, or accuracy.

In summary, most of the studies of non-taped repeated reading indicate that treatment effects are limited to testing conditions that are similar to those used

during training. For the most part, gains on transfer tasks have been limited to measures of fluency.

### *Methodological Limitations*

There is considerable variation in how the fluency building studies were designed and conducted. For example, many studies of repeated reading (e.g., Swain & Allinder, 1996; Turpie & Paratore, 1995; Weinstein & Cooke, 1992), and most studies of previewing (e.g., Rose, 1984a, 1984b, 1984c; Rose & Beattie, 1986; Rose & Sherry, 1984; Sindelar, 1987) used single subject designs. These studies are therefore not appropriate for conventional meta-analysis. With commercially produced reading programs, the samples were usually convenience samples, and the researchers did not have much control as to how accurately the programs were implemented in different schools (e.g., Denton, 1997; Vollands, Topping, & Evans, 1999). Very often, subjects were not randomly assigned to conditions (e.g., Rasinski, 1990), with the result that treatment effects are questionable because it is unclear whether the subjects in the contrasting conditions were comparable before treatment. Further, many studies did not include a control group (e.g., Rasinski, 1990; Dowhower, 1987; Sindelar, Monda, & O'Shea, 1990). Without a control group, it is difficult to determine whether the interventions were causally connected to student performance. In other studies (e.g., Rashotte & Torgesen, 1985), the measures used were not reliable, resulting in an absence of improvement in student performance.

In summary the intervention studies varied considerably by their nature, characteristics, design, methods, and measurement. They have also varied

considerably in their effects on student reading fluency and comprehension. To gain a more accurate picture of these effects, I included studies in this meta-analysis that employ a true or quasi-experimental design. In other words, I included studies that employed random assignment of subjects to conditions, or studies in which the pretreatment equivalence of groups was examined when random assignment is not made. Most importantly, I included only studies that had a control group.

## CHAPTER III

### METHOD

#### Eligibility Criteria for Study Inclusion

Studies for this meta-analysis were selected based on a set of detailed criteria. These criteria follow:

- 1) There must be some intervention or treatment to increase reading fluency. Reading fluency includes reading rate and reading accuracy (see definitions below).
- 2) The intervention must be conducted, at least in part, in public or private elementary school settings. A school may be public or private. Preschools were not eligible. Interventions that included family components were eligible provided that some portion of the program was implemented in a school, or through a school. After-school programs were eligible as long as they were implemented, at least in part, in a school setting, or implemented through a school.
- 3) Intervention effectiveness must be measured in terms of one of the reading fluency indices, i.e., reading rate and reading accuracy. Studies that measured reading fluency and reading comprehension were eligible, but those that measured only reading comprehension were not eligible. Reading rate could have been measured as words read per minute or number of minutes used to read a passage of certain number of words. Reading rate

could also have been measured in reading age or grade equivalent. Reading accuracy could also have been measured as number or percent of errors made, or number of words read correctly per minute, or reading age or grade equivalent. Reading comprehension could have been the number of questions answered correctly, number of propositions recalled, or performance on a standardized reading comprehension test.

- 4) Quantitative data must be reported for reading rate or reading accuracy or both. The study had to report means, standard deviations, results of statistical tests, *p* values, etc. from which an effect size could be calculated for a reading rate and reading accuracy related outcome variable.
- 5) Studies must assess intervention effects on subjects who were elementary-aged children with or without disabilities. Elementary-aged children referred to those attending classes in grades K-6. If elementary-age and older children were included in a study, the study was acceptable only if the results for the “elementary-age” group were reported separately, or if the elementary-age children constituted a majority of the subjects for whom results were reported, and it was reported or it was plausible that the oldest subjects in the sample were not older than 14. Subjects with disabilities included those labeled LD, ED/EBD, mild MR, or speech/language who were receiving special education services. Subjects without disabilities included both students who had no reading problems and those who were considered below-average, non-fluent readers.

- 6) Studies are required to use experimental or quasi-experimental designs that compared subject groups receiving one or more identifiable treatments with one or more control conditions. Randomized or matched post-only study designs with a control group were eligible. Control conditions may have been “no treatment,” “instruction as usual,” placebo treatment, or any other similar condition set up as a contrast to the treatment condition that did not represent a concerted effort to produce change.
- 7) To be eligible as an experimental or quasi-experimental design, a study was required to meet at least one of the following criteria:
  - a) Subjects were assigned randomly to treatment and control conditions.
  - b) If subjects were not randomly assigned to treatment and control conditions, they were matched on pretest variables and/or other relevant personal and demographic characteristics.
  - c) If subjects were not randomly assigned or matched, information about initial group differences on key dependent variables had to be presented in sufficient detail to permit coding of the degree of initial (pretreatment) equivalence. That is, the study must have had both pretests and posttests on one or more eligible outcome variables.
- 8) Post-only non-equivalent comparisons (not randomized or matched) are not eligible. The one-group pretest-posttest design in which measures of one or more of the eligible reading outcomes were taken before treatment and after treatment on the same group of subjects was not eligible. Multiple-group, multiple treatment pre-post studies where more than one subject group was



involved, each receiving a different treatment (but without a control group) were not eligible. Similarly, one-group treatment-control designs that used an ABAB format (single subject design) were not eligible.

- 9) Studies must be conducted in an English speaking country and reported in English. The date of publication was from 1966 to 2001 (Appendix A).

### Retrieval of Research Reports

To obtain eligible studies, Educational Resources Information Center (ERIC), PsycINFO, Exceptional Children, Education Index, Education Abstracts FTX, Dissertation Abstracts International, and International Education were scanned systematically from 1966 to 2001. There were three categories of key words: treatment, population/subjects, and dependent measures. Key words were identified by examining existing literature in reading fluency research and by referring to the thesaurus for computerized databases. Key words for “treatment” included specific interventions such as repeated reading, neurological impress method (NIM), previewing, oral recitation lesson (ORL), class-wide peer tutoring (CWPT), peer tutoring, tutoring, fluency building lessons (FDL), computer-assisted programs, decoding training, and fluency building strategies. Key words for “subjects/population” included elementary, elementary-age, students with disabilities, disabled, poor reader(s), remedial reader(s), nonfluent/disfluent readers, and reading difficulty/problems. Key words for “dependent measures” included reading fluency (speed and/or accuracy), and (reading) comprehension. The abstracts obtained from the searches were examined prior to study selection to

eliminate those that clearly did not meet inclusion criteria (e.g., abstracts of literature reviews). Potentially eligible studies were obtained in libraries on campus, or via interlibrary loan. The reference lists of previous literature reviews and obtained studies were also reviewed to search for eligible studies. Additionally, a hand-search was conducted of all unbound journals that were published in 2001 for possible recent studies on reading fluency. These journals included Educational Research Quarterly, Exceptional Children, Journal of Educational Psychology, Journal of Educational Research, Journal of Learning Disabilities, Journal of Reading Behavior, Journal of Remedial and Special Education, Journal of Special Education, Learning Disability Quarterly, Learning Disability Research, Learning Disability Research & Practice, Learning and Instruction, Reading Improvement, Reading Research and Instruction, Reading Research Quarterly, and The Reading Teacher.

### Coding of Research Reports

Thirty-nine eligible studies were identified and located through the above search process. All eligible studies were coded using a detailed coding scheme based on a coding manual I developed (see Appendices B and C). The coding manual was refined through an iterative process in which issues raised in reviews, and new information from primary studies obtained in the process of preliminary coding was used to add, remove, or clarify variables of interest. For example, in the early stage of coding, it was not clear whether it was important to code how the treatment was delivered (e.g., adult directed, peer mediated, small group, individually), but as the coding process proceeded, it was clear that the format in

which the intervention was carried out was important. So a category was created for it. Specifically, I asked the question, “How was the intervention conducted?” And the options were (a) adult-directed, whole class or small group, (b) adult-directed, one to one, (c) tutoring, peer, (d) tutoring, cross-age, and (e) individual, subjects independently completed treatment tasks.

Two general categories of information were coded: study descriptors (see Appendix D) and effect size information (see Appendix E). Study descriptors were methodological variables (e.g., random assignment, sample size), study context (e.g., year, type of publication), treatment characteristics (e.g., treatment duration, intensity, focus), subject characteristics (e.g., disability, age, sex, ethnicity), and training materials characteristics (e.g., difficulty level, whether they were connected texts or words). Effect size information included means, standard deviations, and other statistics such as *t*-values or *F*-values.

### *Effect Size Coding*

The primary dependent variable in this meta-analysis was the standardized mean difference effect size coded for each treatment-control contrast in experimental and quasi-experimental studies. To compute treatment-control effect size, reading outcomes within each study were identified that reflected differences between the treatment group and control group. Outcomes were divided into those representing reading rate/speed, reading accuracy, and reading comprehension. An effect size estimate was computed for each relevant outcome for which sufficient quantitative information was reported. The index used for this purpose was the

standardized mean difference score, defined as the difference between the posttest mean of the treatment group and the posttest mean of the control group divided by the standard deviation pooled across the treatment and control group (Cohen, 1988; Lipsey & Wilson, 2000).

Mean difference effect sizes were calculated separately for each treatment vs. control comparison and for each outcome of interest in studies that involved only one treatment and one control group, as well as in studies that involved multiple treatment groups. Thus, some studies may have contributed a number of effect sizes to the database because of multiple treatment groups and multiple outcome measures. However, the standardized mean difference effect size index has been shown to be inflated when based on small sample sizes, particularly sample sizes smaller than 20 (Hedges, 1981). Hedges provided a simple procedure to correct this bias and all effect size computations in this meta-analysis were corrected in accordance with the following formula:

$$ES'_{sm} = [1 - 3/(4N - 9)]ES_{sm}$$

In the formula,  $ES'_{sm}$  is the corrected standardized mean difference effect size between the treatment and control group.  $N$  is the total sample size, and  $ES_{sm}$  is the biased standardized mean difference between the treatment and control group.

When means and standard deviations were not available, standardized mean difference effect sizes were estimated from other statistics, such as  $t$ -values and  $F$ -values (see Lipsey & Wilson, 2000).

### *Coding of Study Descriptors*

In addition to effect size information, it is often important to examine relationships between effects and study characteristics. Thus, the following additional study variables were coded.

#### *Methodological Variables*

Information pertaining to the study design was coded, including the experimental design (i.e., type of design, unit and method of assignment to treatment and control conditions, pre-intervention equivalence of treatment and control groups), dependent measures, sample sizes at pretest and posttest, and attrition of subjects. (See Appendix B for details related to this variable and all subsequent study characteristics.)

#### *Study Context*

This category included the year and type of publication (e.g., book, journal article or book chapter, thesis or doctoral dissertation, technical report, or conference paper).

#### *Treatments*

Specific information about the nature of the intervention was coded, including the specific treatment, and its duration, frequency, intensity, and type of grouping (e.g., dyadic, group, or individual). Further, because the interventions covered in the studies for the present meta-analysis were very varied, they were examined in three different ways to reflect different aspects of the treatments in these studies, i.e., specific fluency strategies, how the interventions were conducted (repetitive or non-repetitive), and the type of text involved (connected or non-connected).

Despite ambiguity and difficulty, the interventions first were coded to reflect specific fluency building strategies in five categories: 1) repeated reading and other guided oral reading (of connected texts), 2) repeated reading and other guided oral reading of mixed texts (i.e., with both words and connected texts), 3) encouraging students to read more, 4) general teaching strategies or programs, and 5) word level training.

Then, treatments were coded to based on whether they focused on (a) repetitive practice of sounds, words or phrases; (b) repetitive practice of stories/passages; (c) repetitive practice of sounds, words, phrases, and stories/passages; (d) non-repetitive practice of sounds, words, or phrases; (e) non-repetitive practice of stories/passages; (f) non-repetitive practice of sounds, words, phrases, and stories/passages; and (g) other, including specific instructional methods that encompassed multiple components. These included strategies such as reciprocal teaching, explicit teaching, effective teaching, direct instruction, and commercially produced program (i.e., Reading Recovery). Lastly, treatments were coded based on the type of text involved, reflecting whether the treatment involved non-connected texts (i.e., sounds, words and phrases), or connected text (i.e., stories), or a combination of both connected text and non-connected texts.

The treatment information was coded to help explore characteristics of different treatments that were associated with most and least effective fluency-building practice.

### *Participants*

Where possible, the age, sex, ethnicity, grade level, reading level prior to

treatment, and disability type of the subjects were coded to examine if treatment differed as a function of participant characteristics.

### *Training Materials*

This category of information included the difficulty level of the material, the type of material (list of words or story passages), and whether it was researcher-made or commercially produced. This information was coded to explore if material characteristics were associated with different treatment effects, because some researchers have failed to find improvement in student performance due to a high degree of difficulty of the training materials used (Fleisher, Jenkins, & Pany, 1979).

### *Coding Reliability*

To ensure that the coding of the studies was accurate and consistent, I first coded all eligible reports using my detailed coding scheme. Then, a second coder and I discussed the rationale and content of the meta-analysis, and the role and importance of a second coder. We then discussed the coding manual and coding forms, and we agreed that should there be any questions about the studies, the coder should ask me for clarification. Then, without consulting my first coding, this coder coded 10% of the reports using the same coding manual and coding forms. When there were differences between the two coding passes, they were reconciled by reviewing the report further to select the better alternative. On average, the coding reliability was 85% or above.

## Data Analysis

Data analysis in this meta-analysis involved a sequence of steps. The first was descriptive. The research base was described with regard to research design, student characteristics, treatment characteristics, outcome constructs and the like. Next, the statistical framework and data analysis involved procedures for adjusting posttest effect sizes for pretest differences, adjusting outlying effect sizes (Winsorizing), correcting effect sizes for small sample bias, weighting each effect size by the inverse of its sampling error variance, and testing effect size heterogeneity to guide further analyses. Further, in the present meta-analysis, since studies with multiple treatment groups were included, the effect sizes obtained from the same study shared the same control group, creating interdependencies between these effect sizes. To correct these interdependencies, a procedure proposed by Gleser and Olkin (1994) was used, which is discussed in detail in Chapter IV.

The homogeneity test, or Q statistic (Hedges & Olkin, 1985), was used to explore whether the observed variability in the distribution of effect size estimates was greater than would be expected from sampling error. If the Q statistic was significant, further analyses were warranted to identify the sources of this observed variability in the effect size distribution. The basic procedure in this case was to specify a model (i.e., one or more study descriptor variables) that would successfully divide a heterogeneous group of studies into smaller groups that were homogeneous. There were two general methods to specify the model (see, for example, Hedges, 1994; Raudenbush, 1994). The first, called the fixed effects model, assumed that the effect sizes used in the model came from a sample of



treatment and comparison groups that estimated a true (fixed) effect size (the population parameter). In this case, the individual effect sizes differed from the true value only because of sampling error. In other words, each group included a sample of subjects from the population, and the effect size estimate derived from these subjects would therefore differ from the true population value. In a model based on the fixed effects approach, the between study differences would be due only to the set of study descriptors specified in the model and the subject-level sampling error resulting from the use of sample statistics.

Consider, however, a situation in which the true effect size could itself vary. This variation might have been caused by numerous and unidentifiable variables. In this case the true population effect size would be viewed to be random. Under the random effects model, variability between effect sizes could be due to the set of study descriptors in the model, subject-level sampling error, and other unidentifiable between-study sources of random variation. A model that includes all three of these sources of variability – study descriptors, sampling error, and random variation – is called a mixed effects model because it includes both random and fixed components. Mixed models are more conservative than fixed effects models because they assume that some of the variability among effect sizes is random and unpredictable. For this meta-analysis, the mixed effects model was used and the random effects variance component was estimated using the method of moments procedure (Raudenbush, 1994) and SPSS macros developed by David Wilson (Lipsey & Wilson, 2000).

Once the model was specified, I could examine the various relationships between the set of study predictors in the model, such as subject age or treatment type of duration, and the effect size. I could also draw comparisons across the different types of treatment interventions represented in the database. In this meta-analysis, because studies with multiple treatment groups were included and some studies may have contributed several effect sizes calculated from the same control group, Gleser and Olkin's (1994) multivariate modeling was used to control for interdependence between effect sizes. As mentioned earlier, procedures for correcting the interdependencies among effect sizes obtained from multi-treatment studies are discussed in detail in Chapter IV.

## CHAPTER IV

### RESULTS

The current meta-analysis sought to answer the following questions:

1. What is the average effect size representing each intervention for each reading outcome? Is the average effect size representing each intervention for each outcome significantly different from zero?
2. What is the average effect size representing each intervention for each outcome for students with disabilities? Is the average effect size of each intervention for each outcome for students with disabilities significantly different from zero?
3. What is the average effect size representing each intervention for each outcome for students without disabilities? Is the average effect size of each intervention for each outcome for students without disabilities significantly different from zero?
4. Is the variability in the distribution of effect sizes representing each intervention for each outcome heterogeneous? If the variability in the effect size distribution representing the interventions for each outcome is heterogeneous, what variables are associated with that variability? In other words, what factors are associated with or best predict the least or most effective intervention(s) to increase reading fluency?

Before I respond to these questions, I will provide a profile of the studies in the meta-analysis with regard to experimental design, student characteristics, treatment characteristics, and so forth. The variables described in this section not only serve

as the independent variables for the later analyses, but also help to characterize the literature. This profile describes the types of interventions available and common methodologies, and helps detect gaps in the literature. Next, I present the results of the meta-analysis, focusing primarily on identifying the method, subject, and treatment characteristics that are associated with variability in reading rate, accuracy, and comprehension, and on pinpointing the most effective interventions for improving these three outcome constructs.

### Descriptive Analysis

The search of computer databases from 1966 to 2001 and previous literature reviews yielded 119 study reports. The hand search yielded four more reports on fluency promoting interventions, resulting in a total of 123 reports that were potentially eligible for the meta-analysis. A study report in this meta-analysis is defined as an independent report that has been published 1) in a peer reviewed journal or book, 2) as an ERIC technical report, or 3) as a dissertation. A study report may include results for one sample in which one experimental condition was contrasted against one control or multiple experimental conditions were contrasted against one common control. It may also represent results for several samples. For example, an ERIC document reported results for 1<sup>st</sup> graders, 2<sup>nd</sup> graders and 3<sup>rd</sup> graders separately, each grade level with its own control group. In this case, there would be three independent samples and effect sizes would be calculated for the three samples separately. If it involved only one treatment, this one study report would contribute three treatment-control comparisons to the meta-analysis, one for

each sample or grade level. Using the eligibility criteria already described, I further screened the 123 reports. Eighty-four study reports were excluded because of one of the following reasons:

1. Studies used an ABAB single subject design (n=18).
2. Studies did not have a control group even though they used a group-design (n=30).
3. Studies did not measure reading fluency (n=19).
4. Studies provided inadequate quantitative data to compute effect sizes (n=8).
5. Authors used a repeated measures design where the same subjects served as both treatment subjects and controls (n=4).
6. Subjects were not elementary school students (n=5) (see Appendix F for a list of excluded studies).

The final pool included 39 study reports. Among them, 25 involved single treatments. Of the 25, three included results for two independent samples, and each independent sample had its own control group (Eldredge et al., 1996, normal readers and low readers; Mathes et al., 1998, normal readers and low readers; Vollandts et al., 1999, two separate cohorts of students). Therefore, these three study reports contributed six treatment-control comparisons. One of the 25 study reports represented results for three independent samples, and each sample had its own control (Rashotte et al., 2001, 1st and 2nd grade students, 3rd and 4th students and 5th and 6th grade students). This one report contributed three treatment-control comparisons. The remaining 21 study reports each contributed one treatment-control comparison. Altogether, therefore, there are 30 (21 + 6 + 3) single treatment-

control comparisons. In other words, 30 independent samples obtained from 25 study reports involved single treatments (see Table 1).

Table 1. Number of Samples -Single Treatment Studies

Author(s)	Number of Samples	Number of Comparisons
Allor, Fuchs, & Mathes (2001)	1	1
Aslett (1990)	1	1
Baker, Gersten, & Keating (2000)	1	1
Collins (1994)	1	1
Conte & Humphreys (1989)	1	1
Denton (1997)	1	1
Eldredge, Reutzal, & Hollingsworth (1996)	2	2
Fuchs, Fuchs, & Mathes (1997)	1	1
Gunn, Biglan, Smolkowski, & Ary (2000)	1	1
Hannah (1994)	1	1
Hollingsworth (1970)	1	1
Hunt (1994)	1	1
Jones, Torgesen, & Sexton (1987)	1	1
King (1982)	1	1
Koch (1984)	1	1
Koskinen & Blum (1984)	1	1
Laffey, Kelly, & Perry (1979)	1	1
Mathes, Howard, Allen & Fuchs (1998)	2	2
Mathes & Babyak (2001)	1	1
Miller, Robson & Bushell (1986)	1	1
Rashotte, MacPhee & Torgesen (2001)	3	3
Rasinski, Padak, Linek & Sturtevant (1994)	1	1
Reutzal & Hollingsworth (1993)	1	1
Richek & McTagu (1988)	1	1
Vollands, Topping & Evan (1999)	2	2
Total		30

Ten of the 39 study reports involved two treatments. Seven of these 10 study reports each had one independent sample comparing two treatments against a common control (Arnold, 1972; Cohen, 1989; Fleisher, Jenkins & Pany, 1979; Mathes & Fuchs, 1993; Mathes & Babyak, 2001; Shany & Biemiller, 1995; Simmons

et al., 1995). Each report contributed two treatment-control comparisons (for a total of 14). Three of the 10 study reports each presented results for three independent samples (Fuchs, Fuchs, et al., 2001; normal readers, poor readers and special education students in two treatment conditions: PALS, PALS + Fluency; Fuchs, Fuchs, et al., 2000, normal readers, poor readers and special education students in two treatment conditions: PALS, PALS + Fluency; Fuchs, Fuchs, et al., 1999, normal readers, poor readers and special education students in two treatment conditions: PALS, PALS + Comprehension). Each sample had its own control group and involved two treatments and, as a result, each sample contributed two treatment-control comparisons. In total, these 10 study reports contributed 32 treatment-control comparisons (see Table 2).

One study report (Sindelar, 1982) represented results for three treatments, providing three treatment-control comparisons. Another study presented results for three independent samples (Fuchs, Fuchs, et al., 2000, normal readers, poor readers, and special education students in three treatment conditions: PALS + Decoding + Sound Play, PALS + Decoding, PALS + Sound Play). Each sample had its own control group and involved three treatments. Therefore, with each sample contributing three comparisons, this study report contributed a total of 9 treatment-control comparisons. These two study reports therefore contributed 12 treatment-control comparisons. One study report (Simmons et al., 1994) involved four treatments, yielding 4 treatment-control comparisons. One study report (Marston, Deno & Kim, 1995) presented results for six treatments, producing six treatment-control comparisons (see Table 2).

Table 2. Number of Samples - Multiple Treatment Studies

Author(s)	Number of Samples	Number of Comparisons
<u>Two Treatments</u>		
Arnold (1972)	1	2
Cohen (1989)	1	2
Fleisher, Jenkins & Pany (1979)	1	2
Fuchs et al. (1999)	3	6
Fuchs et al. (2001)	3	6
Fuchs et al. (2000)	3	6
Mathes& Fuchs (1993)	1	2
Mathes & Babyak (2001)	1	2
Shany & Biemiller (1995)	1	2
Simmons, Fuchs, et al. (1995)	1	2
<b>Subtotal</b>		<b>32</b>
<u>Three Treatments</u>		
Fuchs, Fuchs et al. (2000)	3	9
Sindelar (1982)	1	3
<b>Subtotal</b>		<b>12</b>
<u>Four treatments</u>		
Simmons et al. (1994)	1	4
<b>Subtotal</b>		<b>4</b>
<u>Six treatments</u>		
Marston, Deno & Kim (1995)	1	6
<b>Subtotal</b>		<b>6</b>
<b>Grand Total</b>		<b>54</b>

The final pool for the present meta-analysis consisted of 84 treatment-control comparisons (each comparison was like a case or subject in SPSS), obtained from 39 study reports. Of the 84 comparisons, 30 were from single treatment studies, 32 from two-treatment studies, 12 from three-treatment studies, 4 from a four-treatment study, and 6 from one six-treatment study (see Tables 1 & 2). Over half of the comparisons were published in journals or books (n=45), but unpublished dissertations, theses, and technical reports (n=39) were well represented in the



sample. Table 3 presents summaries of treatment characteristics in the meta-analysis. Table 4 summarizes the general characteristics of the subjects.

Table 3. Characteristics of Treatments

Variables	N <sup>a</sup>	%
Sample size		
10-30	29	34.5
31-50	25	29.8
51-100	14	16.7
>100	16	19.0
Study design		
Random assignment	77	91.7
Quasi-experimental, nonrandom	7	8.3
Treatment duration (in weeks)		
1 to 6 weeks	13	15.5
7 to 20 weeks	47	56.0
21 to 36 weeks	24	28.6
Frequency of treatment		
1-2 times a week	4	4.8
3-4 times a week	53	63.1
Daily	27	32.1
Treatment length (in minutes)		
15-30 minutes	36	42.9
31-45 minutes	44	52.4
46-60 minutes	4	4.8
Treatment focus		
Fluency only	44	52.4
Comprehension only	9	10.7
Fluency and comprehension	31	36.9
Treatment format		
Teacher-directed, small group or whole class	21	25.0
Tutoring, adults, peers or cross-age	54	64.3
Subject independently, with supervision	9	10.7

<sup>a</sup> Number of treatment-control comparisons.

Table 4. Characteristics of Subjects

Variable	N <sup>a</sup>	%
Age		
5-8 years	34	40.5
9-12 years	15	17.9
Unknown (not reported)	35	41.7
Disability/remedial		
Students with disabilities	26	31.0
Remedial readers	31	36.9
Normal readers (average or above average)	20	23.8
Mixed readers (of all three types)	7	8.3
Ethnicity		
>60% white	10	11.9
>60% black	3	3.6
>60% Hispanic	1	1.2
Mixed race (none more than 60%)	35	41.7
Unknown (not reported)	35	41.7
Grade level		
Grades K-2	46	54.8
Grades 3-4	12	14.3
Grades 5-6	7	8.3
All grade levels	19	22.6
Reading level		
Read at grade level	20	23.8
1 year below grade level	43	51.2
2 years below grade level	8	9.5
3 years below grade level	3	3.6
Unknown (not reported)	10	11.9
Sex		
5%-49% male	7	8.3
50% male	3	3.6
51%-95% male	43	51.2
Unknown (not reported)	31	36.9

<sup>a</sup> Number of treatment-control comparisons.

Since the focus of this meta-analysis included three outcome constructs, i.e., reading rate, accuracy and comprehension, each treatment-control comparison could potentially contribute three effect sizes, one for each of the three outcome variables. However, some reports did not measure all three outcomes. As a result, the 84 treatment-control comparisons yielded 162 effect sizes, 75 for reading rate, 32 for reading accuracy, and 55 for reading comprehension (see Table 5 for the distribution of effect sizes in relation to number of treatments).

Table 5. Distribution of Effect Sizes in Relation to Number of Treatments

# of Treatments/Samples	# of Reports	# of ES Contributed for			Total # of ES
		Rate	Accuracy	Comp.	
<b>Single Treatment</b>					
One Sample	21	16	13	15	44
Two Samples	3	4	4	6	14
Three Samples	1	3	3	3	9
Sub Total	<b>25</b>	<b>23</b>	<b>20</b>	<b>24</b>	<b>67</b>
<b>Two Treatments</b>					
One Sample	7	12	6	12	30
Three Samples	3	18	6	12	36
Sub Total	<b>10</b>	<b>30</b>	<b>12</b>	<b>24</b>	<b>66</b>
<b>Three Treatments</b>					
One Sample	1	3	0	3	6
Three Samples	1	9	0	0	9
Sub Total	<b>2</b>	<b>12</b>	<b>0</b>	<b>3</b>	<b>15</b>
<b>Four Treatments</b>					
One Sample	1	4	0	4	8
Sub Total	<b>1</b>	<b>4</b>	<b>0</b>	<b>4</b>	<b>8</b>
<b>Six Treatments</b>					
One Sample	1	6	0	0	6
Sub Total	<b>1</b>	<b>6</b>	<b>0</b>	<b>0</b>	<b>6</b>
<b>Column Total</b>	<b>39</b>	<b>75</b>	<b>32</b>	<b>55</b>	<b>162</b>

ES: Effect size.

Comp.: Comprehension.

## Effect Size Adjustments

### *Adjustments of All Effect Sizes*

Before discussing the important results, it is necessary to describe the procedures used to adjust effect sizes in this meta-analysis. Four procedures, i.e., pretest effect size adjustment, Winsorizing outliers in the effect size distribution, correcting for small sample size bias, and weighting effect sizes by the inverse of the sampling error variance, were applied to the adjustments of the effect sizes. In the following section, I describe each of the four procedures.

The first procedure involved adjusting posttest effect sizes for pretest differences. The focus of the present meta-analysis is posttest effect sizes, but some studies provided both pretest and posttest scores, therefore, both pretest and posttest effect sizes were computed for these studies. Unfortunately, there were not enough pretest effect sizes in each outcome category to warrant systematic analysis. However, some pretest effect sizes were obviously large, indicating there were cases of serious pretreatment differences between treatment and control groups and ignoring them would skew the distribution of posttest effect sizes. I therefore decided to adjust posttest effect sizes by subtracting pretest effect sizes from posttest effect sizes to get rid of pretreatment differences as much as I could. This adjustment was done for all studies that had a pretest effect size.

Specifically, for reading rate, there were 35 (or 47%) pretest effect sizes all of which came from experimental design studies. There were two more comparisons (from nonrandomized studies) for which this adjustment would be particularly

appropriate (because of non-randomized designs) but no pretest effect sizes were provided. For reading accuracy, there were 13 (or 41%) pretest effect sizes, and three were from non-experimental design studies. There were four more comparisons (from nonrandomized studies) for which this adjustment would have been appropriate but no pretest effect sizes were provided. For reading comprehension, there were 28 (or 51%) pretest effect sizes, and three of them were from nonrandomized studies. There were three more effect sizes (from nonrandomized studies) for which this adjustment would have been appropriate but no pretest effect sizes were provided. In general, though, there were only a small number of comparisons from nonrandomized studies, i.e., there were two, seven and six comparisons from nonrandomized studies for reading rate, accuracy and comprehension, respectively.

The second procedure involved examining extreme values in the distributions of effect sizes for all three outcome constructs. Extreme effect size values may not accurately reflect the body of research that a meta-analysis seeks to summarize and may, in fact, distort the results of statistical analyses (Hedges & Olkin, 1985). Several large effect sizes were found in the effect size distribution of each outcome category. A careful examination of the shapes of the distributions suggested that these were extreme values. The studies that generated these outliers were examined carefully to make sure that the coding was accurate. These outliers could seriously skew the effect size distribution, however, I was reluctant to remove them from analyses because there were so few studies and removing them may lead to an underestimate of the effect of the specific type of intervention that effect size

represented. I decided to recode these extreme values to more moderate values through a process called Winsorizing (Lipsey & Wilson, 2000), resetting these extreme values to be just outside the next largest non-outlier values and making the difference between these outliers and the rest of the effect sizes in the distribution less extreme to fit the distribution better. However, these outliers were still on the end of the distribution and provided the biggest values in the distribution, therefore, their relative importance was maintained. Specifically, for reading rate, there were three outlying effect sizes, -.48, 1.97 and 3.14, and they were recoded to -.25, 1.34 and 1.40. For reading accuracy, there were two outliers, 1.79, and 4.92. These two outliers were recoded to 1.40, and 1.55, respectively. For reading comprehension, there were four outlying effect sizes, -.64, 1.29, 1.45, and 2.37, which were recoded to -.50, 1.10, 1.20, and 1.35, respectively.

Because each effect size would be weighted by the inverse of its variance and the inverse of the variance of an effect size is strongly influenced by its sample size (more about this below), effect sizes with extremely large sample sizes would be weighted heavily in an analysis. Therefore, the sample sizes were also examined carefully and there were no extreme sample sizes.

The third procedure involved correcting for small sample bias. In general, effect sizes based on small samples are biased estimates of effect. This bias was removed by using the following formula (Hedges & Olkin, 1985; Lipsey & Wilson, 2000):  $[1 - 3/(4n-9)]*d$ , where  $n$  is the total sample size (treatment  $n$  + control  $n$ ), and  $d$  is the unadjusted or biased effect size. All analyses in this meta-analysis used these  $n$ -adjusted effect sizes.

The fourth procedure dealt with weighting each effect size by the inverse of its sampling error variance. Since effect sizes are derived from sample statistics (e.g., means and standard deviations), those derived from large samples contain less sampling error than those obtained from small samples. Therefore, when combining effect sizes for statistical analysis, each effect size is weighted by the inverse of the sampling error variance ( $w_i$ ) so that its contribution was proportionate to its reliability (Hedges & Olkin, 1985). The computational formula for estimating the sampling error variance for standardized mean difference effect sizes was given by Hedges and Olkin (1985) as follows:

$$\text{var}(d_i) = (n_i^t + n_i^c) / n_i^t n_i^c + (d_i)^2 / 2(n_i^t + n_i^c)$$

where  $n_i^t$  is the treatment group sample size for the  $i$ th study,  $n_i^c$  is the control group sample size for the  $i$ th study, and  $d_i$  is the  $n$ -adjusted standardized mean difference effect size for study  $i$ . This formula works for effect sizes derived from studies with only one treatment group and one control group. However, in the present meta-analysis, some studies with multiple treatment groups were included, and the effect sizes obtained from the same study shared the same control group, creating interdependencies between these effect sizes. To correct these interdependencies, I used a procedure proposed by Gleser and Olkin (1994), which is discussed in the following section.<sup>1</sup>

#### *Independence among Effect Sizes from Multiple Treatment Studies*

In order to meet the assumptions for later statistical analyses in a meta-analysis, the effect sizes in a distribution must be statistically independent. If for

---

<sup>1</sup> Dr. Mark Lipsey provided valuable advice on the procedure.

each study,  $i$ , there were only one effect size,  $d_i$ , it would be safe to assume that the effect sizes were statistically independent. Subsequently, it would be safe to run all analyses on these effect sizes using the appropriate inverse variance weight ( $w_i$ ) described above.

Consider, however, a situation in which one study,  $i$ , because of multiple treatments compared with the same control group, produced multiple effect sizes for the same outcome construct,  $d_{ij}$ , with  $j = 1, 2, \dots$ . To maintain statistical independence of the effect sizes in the distribution, two types of weighting are needed on each effect size ( $d_{ij}$ ) obtained from the study. One is the inverse variance weight ( $w_i$ ) used for effect sizes from single treatment studies, as described previously. The other is the weight that deals with the inter-dependencies among the multiple effect sizes obtained from studies with more than one treatment. In other words, we not only need to address the appropriate proportionate representation of each effect size so that the multiple effect sizes from one particular study still contribute in essence the net effect of one effect size to the analysis, we also need to address the problem of overlap (i.e., inter-dependencies) between the effect sizes because of the common control group. To account for this overlap, weights are obtained from the variance-covariance matrix describing the dependencies caused by the sharing of a common control group. The following description will help readers understand how the procedure works.

Normally, when a study with one treatment contributes more than one effect size for the same outcome, we would average the effect sizes so that there was just one effect size per study for any given analysis. However, in the case of a study with



multiple treatments, rather than combining the individual effect sizes in an average, we could instead maintain them by retaining the weight on each effect size so that while the study contributes more than one effect size, each effect size is weighted in such a way that the multiple effect sizes would be represented according to their respective proportion, and would therefore contribute the net effect of only one effect size. For example, if a study contributes three effect sizes, these effect sizes would be entered on the analysis as  $w_i d_{i1} 1/3 + w_i d_{i2} 1/3 + w_i d_{i3} 1/3$ . This represents the individual effect sizes from the one study in appropriate proportion but does not address the problem of statistical dependencies among them.

If we represent the above situation more generally, we get two types of weighting on each effect size ( $d_{ij}$ ) obtained from one study - the inverse variance weight ( $w_i$ ) and the weight that allocates relative emphasis among the effect sizes. Following Gleser and Olkin (1994), we will represent this second weight as  $a_{ij}$  and call it an  $a$  weight. The above formula for the three effect size values then becomes  $w_i d_{i1} a_{i1} + w_i d_{i2} a_{i2} + w_i d_{i3} a_{i3}$ , with the mean being the special case of the  $a_{ij}$  values all equaling one divided by the number of effect sizes contributed by the study.

Now we need to compute  $a_{ij}$  weights that will behave like regression coefficients and weight the effect sizes ( $d_{ij}$ ) in such a manner that each effect size is making a unique contribution that is statistically independent of the others from the same study. When that is achieved, the effect sizes can all be included in the same analysis without concern about having multiple effect sizes from the same study or about them being statistically dependent. To achieve this, we first had to compute the variance-covariance matrices for effect sizes derived from studies with multiple

treatment groups. Based on Gleser and Olkin (1994), the following formulas were used:

$$\text{var}(d_{jj}) = 1/n_j + (1+0.5d_j^2)/n_0;$$

$$\text{covar}(d_{jj^*}) = (1 + 0.5d_j d_{j^*})/n_0, j \sim j^*, j = 1, \dots, p.$$

where  $n_j$  is the sample size for treatment  $j$ ;  $d_j$  is the effect size for treatment  $j$ ;  $n_{j^*}$  is the effect size for another treatment,  $j^*$ , in the same study, while  $d_{j^*}$  is the effect size for treatment  $j^*$ , and  $n_0$  is the sample size for the shared control group.

Then, the variance-covariance matrices obtained using the above formulas were inverted to get the inverse matrices. From the inverse matrices, the  $a_i$  weights and respective  $w_i$  inverse variance weights were obtained for each of the effect sizes from multiple treatment studies. These procedures for computing the variance-covariance matrices and inverse matrices were performed through Microsoft Excel, involving the  $n$ -adjusted effect sizes for each outcome construct.

In essence, then, to include in a meta-analysis multiple effect sizes contrasting different treatments against the same control group in the same study, in addition to the usual inverse variance weight, we introduced another weighting function,  $a_{ij}$ , that addressed the statistical dependencies and proportionate representation of multiple effect sizes for studies that contributed multiple effect sizes. Instead of the overall weighted analysis involving a set of inverse variance weighted effect sizes,  $w_i d_i$ , we then had a set of double weighted effect sizes,  $w_i d_i a_{ij}$ . In cases where there is only one effect size per study, the  $a_i$  weight equals 1; with more than one effect size per study, the  $a$  weights sum to 1. Through these procedures, an independent distribution of effect sizes was established. All

subsequent analyses in the present meta-analysis were based on these double weighted effect sizes.

### A Summary of Effect Sizes for Outcomes

This summary answers the following questions. First, what is the average effect size representing all interventions for each reading outcome? Is the average effect size for each outcome significantly different from zero? What is the average effect size representing each intervention for each outcome? Is the average effect size for each outcome significantly different from zero? What is the average effect size for all the interventions for each outcome for different student types? Is the average effect size significantly different from zero? This summary also presents the mean effect sizes and 95% confidence intervals for all three outcomes by different treatment features, such as treatment duration, frequency, focus or intensity. In addition, this section presents the results of the homogeneity tests (or the Q statistics) for treatment types, student type and different treatment features. The Q statistic tests whether the observed variability in the distribution of effect size estimates is greater than would be expected from subject-level sampling error.

### *Overall Mean Effect Sizes for Outcomes*

The 84 treatment-control comparisons yielded 162 effect sizes on the three outcomes, with 75, 32, and 55 for reading rate, reading accuracy, and reading comprehension, respectively. The mean effect sizes and confidence intervals in Table 6 were based on a random effects model. A fixed-effects model is fit on the

data and assumes only one source of variability, i.e., the variability within studies. It also assumes that the student populations across studies are sufficiently similar and that the results are suitable to pool together. A random-effects model, on the other hand, assumes a second source of variability (in addition to within-study variability), i.e., the variability among studies. Variation among studies implies that each study potentially estimates different effect sizes. Random-effects models are more conservative in the sense that they allow for more variability in treatment effects and, correspondingly, have larger standard errors.

The mean effect sizes in Table 6 are the averages of all effect sizes for each outcome across treatment types. As Table 6 shows, the mean effect sizes for reading rate, reading accuracy, and reading comprehension were all positive but were fairly small, ranging from small to medium effect (Cohen, 1988). The 95% confidence intervals did not include zero, indicating that the mean effect sizes were significantly different from zero.

Table 6. Mean Effect Sizes for Outcomes

Outcome	Mean ES	N	95% C. I.		Q <sub>B</sub>	
			Lower	Upper		
Reading Rate	0.30	75	0.22	0.37	79.75	<i>ns</i>
Reading Accuracy	0.41	32	0.29	0.53	33.39	<i>ns</i>
Reading Comprehension	0.33	55	0.24	0.43	67.19	<i>ns</i>

1. ES: Effect size.
2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.
3. Q<sub>B</sub>: from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.

The last column of Table 6 presents the results of the homogeneity tests. The Q statistic tests whether the observed variability in the distribution of effect size estimates is greater than would be expected from subject-level sampling error. The statistically non-significant Q statistics for reading rate, reading accuracy and comprehension outcomes seem to suggest that the effect sizes in the distributions do not differ significantly beyond sampling error; however, in the present meta-analysis, that may not necessarily lead to the conclusion that systematic variation in the distributions of the effect sizes is absent. One possible explanation for the non-significant Q is that Q-test does not have high statistical power when applied to modest numbers of studies, especially studies with small samples. In this meta-analysis, forty (40) study reports contributed 85 treatment-control comparisons (which is the unit of analysis) and two-thirds of these comparisons had total samples equal to or smaller than 50. Specifically, 48 (or 64%) of the 75 cases for reading rate, 26 (or 81%) of the 32 cases for reading accuracy and 39 (or 71%) of the 55 cases for reading comprehension had total samples equal to or smaller than 50. Since the effect sizes in the present meta-analysis were obtained from studies involving both single treatment and multiple treatments, another possible explanation for the lack of statistical significance in the Q-tests may be due to the relatively large number of effect sizes from studies with multiple treatments (see Table 5 for the distribution of effect sizes in different types of studies). In spite of great efforts made to create an independent distribution of effect sizes, it is possible that the effect sizes from the same studies were still fairly similar to each other even after they were weighted to function as independent data points. The similarities within these studies

likely were large enough to reduce the overall variability in the effect size distribution below significance level. Given these factors and the research questions this meta-analysis set out to answer, rather than concluding that there was no systematic variability in the effect size distributions, I decided to cautiously analyze a few moderator variables. These analyses included examining the mean effect sizes of different student types, treatment groups, treatment features, and weighted multiple regressions on the three outcome constructs. Because the non-significant Q-tests indicate that there is little, if any, variability beyond subject-level sampling error, all further analyses are fixed effects analyses. In addition, for all these analyses, I have selected to report statistical significance at  $\alpha = .10$  rather than  $.05$  because of the small sample size and associated low statistical power.

#### *Mean Effect Sizes by Student Type*

To examine whether there were any differences in the treatment effects for different student types, i.e., students with disabilities, remedial readers, normal readers, and groups of mixed students (i.e., samples included students with disabilities, normal readers and remedial readers), the mean effect sizes for the three outcome measures were broken down by student types (see Table 7). The mean effect sizes for all three outcome constructs for all student types were positive and the 95% confidence intervals did not include zero. The Q statistic for reading accuracy was significant, indicating that there was systematic variability in the treatment effect sizes for this outcome construct among the four student types. Treatment involving students with disabilities produced largest effect (.86) and those

involving normal readers produced the smallest effect (.26). The Q statistics for reading rate and comprehension were not significant, which indicates that there was no systematic variability in the treatment effect sizes for these two outcome constructs among the four student types. Interestingly, the pattern was that studies involving students with disabilities and remedial readers seemed to have produced larger effects than those involving normal readers.

Table 7. Mean Effect Sizes for Outcomes by Student Type

Outcome/Student Type	Mean ES	N	95% C. I.		Q <sub>w</sub>	Q <sub>B</sub>
			Lower	Upper		
Reading Rate						5.68
Student with Disabilities	0.34	26	0.14	0.53	14.08	
Remedial Readers	0.37	25	0.26	0.48	37.55	**
Normal Readers	0.18	17	0.06	0.30	20.13	
Mixed	0.26	7	0.02	0.50	2.30	
Reading Accuracy						7.79 *
Student with Disabilities	0.86	6	0.45	1.27	1.77	
Remedial Readers	0.42	16	0.25	0.59	14.74	
Normal Readers	0.26	8	0.08	0.44	8.05	
Mixed	0.57	2	0.12	1.03	1.03	
Reading Comprehension						4.70
Student with Disabilities	0.32	12	0.06	0.58	5.58	
Remedial Readers	0.37	24	0.24	0.50	22.52	
Normal Readers	0.20	12	0.06	0.34	28.49	***
Mixed	0.47	7	0.23	0.71	5.89	

1. ES: Effect size.

2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.

3. Q<sub>B</sub>: from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.

\* p<.10; \*\* p <.05; \*\*\* p = .005

### *Mean Effect Sizes by Treatment Type*

The interventions covered in the studies for the present meta-analysis were very varied, making it difficult to group or categorize them in a manner that

would accurately capture and fully represent the common features of these interventions. Therefore, I tried to code them in three different ways to reflect different aspects of the treatments in these studies, i.e., specific fluency strategies, how the interventions were conducted (repetitive or non-repetitive), and the type of text involved (connected or non-connected texts).

### *Recoding of Treatment Type*

Despite ambiguity and difficulty, I first coded the interventions to reflect specific fluency building strategies in five categories: 1) repeated reading and other guided oral reading (of connected texts), 2) repeated reading and other guided oral reading of mixed texts (i.e., with both words and connected texts), 3) encouraging students to read more, 4) general teaching strategies or programs, and 5) word level training.

Included in the first category were strategies such as repeated reading, echo reading, paired reading, neurological impress method (NIM), assisted reading, listening while reading, Fluency Development Lesson (FDL), Oral Recitation Lesson (ORL), etc. The second category was similar to the first one, but the interventions included repeated reading or other guided oral reading of both connected and non-connected texts, such as Peer Assisted Learning Strategies (PALS). The third category, or encouraging students to read more included strategies such as sustained silent reading (SSR), uninterrupted sustained silent reading (USSR), super quiet reading time (SQUIRT), Drop Everything and Read (DEAR), Accelerated Reading (AR), and other reading activities like free reading or independent reading. Only two strategies -SSR and AR- were represented in the current meta-analysis.



The fourth category covered strategies beyond pure oral reading practice. Instead, they usually examined the effect of a specific teaching method or strategy, such as effective teaching, reciprocal teaching, direct instruction, or a specific program such as SRA, HOLT, Reading Recovery, etc. The fifth category was word level training, including training in word recognition, decoding, etc.

These five categories, however, were far from clear-cut because the strategies were varied. In addition, there were very few studies in the last three categories. Therefore, I decided to go beyond specific fluency building strategies and look at some common features these interventions seemed to share. Hence two more codes were added as follows. First, the interventions were coded based on how they were done, indicating whether they focused on 1) repetitive practice of non-connected texts (i.e., sounds, words and phrases); 2) repetitive practice of connected texts (i.e., stories); 3) repetitive practice of both connected and non-connected texts (i.e., sounds, words, phrases, and stories); and 4) non-repetitive practice of connected and/or non-connected texts (i.e., sounds, words, phrases, and stories). The categories in this group were collapsed because there were only a few studies in the non-repetitive practice category. Then the interventions were coded based on the type of text involved, reflecting whether the treatment involved non-connected texts (i.e., sounds, words and phrases), or connected text (i.e., stories), or a combination of both connected text and non-connected texts.

#### *Mean Effect Sizes by Treatment Type*

The mean effect sizes and confidence intervals were obtained for each category of the three types of codes for treatment. Table 8.1 presents the mean

effect sizes and confidence intervals for different fluency building strategies. As the table shows, both repeated reading and other guided oral reading of connected texts and those of mixed texts consistently produced medium effect for all three outcomes (range from .25 to .47) (Cohen, 1988), and were significantly different from zero. Repeated reading and other guided oral reading of connected texts seem to have an edge over repeated reading and other guided oral reading of mixed texts in both accuracy (.47 vs. .38) and comprehension (.45 vs. .25), but not in reading rate (.36 vs. .32). The other three categories produced similar effects except for word training on accuracy (.90) and encouraging students to read more on comprehension (.34). None of them were significantly different from zero (see Table 8.1).

Due to the small number of studies in each category, I decided to collapse the three categories and obtain the mean effect sizes and confidence intervals for them. Table 8.2 presents the means effect sizes and confidence intervals. In addition, I decided to collapse the two categories in repeated reading and other guided oral reading and obtain the mean effect sizes and confidence intervals (see Table 8.3). As Table 8.3 shows, repeated reading and other guided oral reading were clearly somewhat better than other methods.

Table 8.1. Mean Effect Sizes for Different Fluency Building Strategies (1)

Outcome/text level	Mean ES	N	95% C.I.		Q <sub>w</sub>	Q <sub>B</sub>
			Lower	Upper		
Reading Rate						6.61
RR/guided OR of text	0.36	22	0.21	0.50	16.70	
RR/guided OR of mixed text	0.32	35	0.22	0.41	53.07	**
Encouraging to read more	-0.03	1	-0.78	0.72	0.00	
General strategies/programs	0.07	10	-0.12	0.27	0.58	
Word Level Training	0.21	7	-0.07	1.46	2.78	
Reading Accuracy						3.80
RR/guided OR of text	0.47	15	0.29	0.66	17.50	
RR/guided OR of mixed text	0.38	12	0.22	0.54	11.57	
Encouraging to read more	0.14	2	-0.31	0.60	0.27	
General strategies/programs	0.05	2	-0.73	0.83	0.26	
Word Level Training	0.90	1	-0.02	1.82	0.00	
Reading Comprehension						6.04
RR/guided OR of text	0.45	22	0.31	0.60	24.68	
RR/guided OR of mixed text	0.25	23	0.14	0.36	29.97	
Encouraging to read more	0.34	3	-0.06	0.74	4.35	
General strategies/programs	0.09	4	-0.36	0.53	1.25	
Word Level Training	0.12	3	-0.48	0.72	0.91	

1. ES: Effect size; RR = repeated reading; OR = oral reading.
  2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.
  3. Q<sub>B</sub>: from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.
- \*\* p<.05

Table 8.2. Mean Effect Sizes for Different Fluency Building Strategies (2)

Outcome/text level	Mean ES	N	95% C.I.		Q <sub>w</sub>	Q <sub>B</sub>
			Lower	Upper		
Reading Rate						5.88 *
RR/guided OR of text	0.36	22	0.21	0.50	16.70	
RR/guided OR of mixed text	0.32	35	0.22	0.41	53.07 **	
Other	0.11	18	-0.05	0.27	4.09	
Reading Accuracy						1.41
RR/guided OR of text	0.47	15	0.29	0.66	17.50	
RR/guided OR of mixed text	0.38	12	0.22	0.54	11.57	
Other	0.24	5	-0.12	0.60	2.92	
Reading Comprehension						5.27 *
RR/guided OR of text	0.45	22	0.31	0.60	24.68	
RR/guided OR of mixed text	0.25	23	0.14	0.36	29.97	
Other	0.21	10	-0.06	0.48	7.26	

1. ES: Effect size; RR = repeated reading; OR = oral reading.
  2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.
  3. Q<sub>B</sub>: from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.
- \*\* p<.05

Table 8.3 Mean Effect Sizes for Different Fluency Building Strategies (3)

Outcome/text level	Mean ES	N	95% C.I.		Q <sub>w</sub>	Q <sub>B</sub>
			Lower	Upper		
Reading Rate						5.70 **
RR/guided oral reading	0.33	57	0.25	0.41	69.95 *	
Other	0.11	18	-0.05	0.27	4.09	
Reading Accuracy						0.85
RR/guided oral reading	0.42	15	0.30	0.54	29.63	
Other	0.24	12	-0.12	0.60	2.92	
Reading Comprehension						0.63
RR/guided oral reading	0.32	45	0.24	0.41	59.29 *	
Other	0.21	10	-0.06	0.48	7.26	

1. ES: Effect size; RR = repeated reading; OR = oral reading.
  2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.
  3. Q<sub>B</sub>: from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.
- \*\* p<.05

Tables 9.1 and 9.2 present the mean effect sizes for different practice levels. Table 9.1 shows the mean effect sizes and confidence intervals for each outcome at different reading practice levels. The  $Q_B$  statistics for reading accuracy and reading comprehension, which test the differences between practice levels, were significant ( $Q_B = 7.05$ ,  $p < .10$ ;  $Q_B = 9.25$ ,  $p < .05$ , for accuracy and comprehension, respectively). A closer look revealed that for comprehension, repetitive practice at text level produced the largest effect (.52), followed by repetitive practice at the text and non-text level (.25), with the smallest effect produced by repetitive practice at non-text level (i.e., words) (.12). Non-repetitive practice produced a similar effect (.20). For accuracy, for repetitive practice at word level produced the largest effect was, but that effect was from only one study. Repetitive practice at the text level produced the second largest effect (.57), followed by repetitive practice that involved both words and stories. For reading rate, although the  $Q_B$  statistics were not statistically significant, a similar pattern was observed in the mean effect sizes at different practice levels. In other words, repetitive practice at text level produced the largest effect, followed by repetitive practice at text and non-text level, repetitive practice at word level and non-repetitive practice.

When results in Table 9.1 were collapsed across repetitive practice levels, the  $Q_B$  for reading accuracy and comprehension were still significant, and the results indicated that repetitive practice produced significantly larger effect (.46 and .34) than non-repetitive practice (.14 and .20). The  $Q$  value for reading rate was again not significant, probably due to the low power of the  $Q$  test, but the pattern was very clear (see Table 9.2). Repetitive practice produced moderate mean effect (.30), and

the effects were significantly different from zero (as shown by the 95% confidence intervals, which did not include zero), while the mean effect size for non-repetitive was not significantly different from zero.

Table 9.1 Mean Effect Sizes for Outcomes at Different Practice Levels (1)

Outcome/Practice Level	Mean ES	N	95% C.I.		Q <sub>w</sub>	Q <sub>B</sub>
			Lower	Upper		
Reading Rate						1.92
Nonrepetitive practice	0.15	13	-0.11	0.42	3.00	
Repetitive at nontext level	0.21	7	-0.07	0.49	2.78	
Repetitive at text & nontext level	0.29	38	0.20	0.37	56.27	**
Repetitive at text level	0.35	17	0.19	0.51	15.77	
Reading Accuracy						7.05*
Nonrepetitive practice	0.14	9	-0.13	0.42	0.94	
Repetitive at nontext level	0.90	1	-0.02	1.82	0.00	
Repetitive at text & nontext level	0.38	12	0.22	0.54	11.57	
Repetitive at text level	0.57	10	0.36	0.78	13.84	*
Reading Comprehension						9.25**
Nonrepetitive practice	0.20	14	-0.03	0.42	12.25	
Repetitive at nontext level	0.12	3	-0.48	0.72	0.91	
Repetitive at text & nontext level	0.25	23	0.14	0.36	29.97	
Repetitive at text level	0.52	15	0.36	0.69	14.82	

1. ES: Effect size.

2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.

3. Q<sub>B</sub>: from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.

\* p<.10; \*\* p<.05

Table 9.2. Mean Effect Sizes for Outcomes at Different Practice Levels (2)

Outcome/Practice Level	Mean ES	N	95% C.I.		Q <sub>w</sub>	Q <sub>B</sub>
			Lower	Upper		
Reading Rate						1.02
Nonrepetitive practice	0.15	13	-0.11	0.42	3.00	
Repetitive practice	0.30	62	0.22	0.37	75.71 *	
Reading Accuracy						4.18
Nonrepetitive practice	0.14	9	-0.13	0.42	0.94	
Repetitive practice	0.46	23	0.33	0.58	28.27	
Reading Comprehension						3.66*
Nonrepetitive practice	0.20	14	-0.03	0.42	12.25	
Repetitive practice	0.34	41	0.24	0.42	53.67 *	

1. ES: Effect size.

2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.

3. Q<sub>B</sub>: from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.

\* p<.10; \*\* p<.05

Table 10 presents the mean effect sizes and confidence intervals for treatment at different text levels. As Table 10 shows, most mean effect sizes at the different levels were within the small to medium effect range, but the Q-between values for different text levels for all three outcome constructs were non-significant, suggesting that treatment effects did not differ systematically as a function of the type of text used or involved in the intervention. However, the pattern for reading comprehension suggested that interventions involving stories produced much larger effect than those involving words or mixed texts (.43 vs. .25 and .12).

Table 10. Mean Effect Sizes for Outcomes at Different Text Levels

Outcome/text level	Mean ES	N	95% C.I.		Q <sub>w</sub>	Q <sub>B</sub>
			Lower	Upper		
Reading Rate						0.66
Sounds/words level	0.21	7	-0.07	0.49	2.78	
Sounds/words, and stories	0.28	43	0.19	0.36	57.55 *	
Story level	0.33	25	0.18	0.47	18.76	
Reading Accuracy						1.22
Sounds/words level	0.90	1	-0.02	1.82	0.00	
Sounds/words, and stories	0.38	13	0.22	0.54	11.59	
Story level	0.41	18	0.24	0.58	20.58	
Reading Comprehension						4.48
Sounds/words level	0.12	3	-0.48	0.72	0.91	
Sounds/words, and stories	0.25	26	0.14	0.36	30.59	
Story level	0.43	26	0.29	0.56	31.21	

1. ES: Effect size.

2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.

3. Q<sub>B</sub>: from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.

\* p<.10.

### *Mean Effect Sizes by Treatment Features*

A series of variables was coded to describe treatment features, including treatment duration (1=1-6 weeks, 2=7-20 weeks, 3=21-36 weeks), treatment frequency (1=1-2 times per week, 2=3-4 times per week, 3=daily), treatment format (1= teacher-directed whole class or small group instruction, 2= tutoring - adults, peers or cross age, or 3= students working independently), length of each treatment session (1=15-30 minutes, 2=31-45 minutes, 3=46-60 minutes), treatment focus (1=fluency only, 2=comprehension only, 3= fluency and comprehension), treatment intensity or amount of treatment time per week, which was obtained by multiplying/combining treatment frequency and treatment session length (1 = 0-60



minutes per week, 2 = 61-120 minutes per week, 3= more than 120 minutes per week), and total treatment time, which was obtained by combining/multiplying treatment frequency, treatment session length and treatment duration (1=0-15 hours, 2=15.01-30 hours, 3=30.01-40 hours, 4=more than 40 hours). For each of the above treatment features, the mean effect sizes and confidence intervals for the three outcome measures were obtained to examine whether treatment effects differed as a function of these different features (Tables 11-17).

### *Treatment Duration*

Table 11 presents the mean effect sizes and confidence intervals for each outcome at different treatment duration levels. The mean effect sizes for all three outcome constructs for all treatment duration levels were positive and the 95% confidence intervals did not include zero. Further, for all three outcome constructs, the  $Q_B$  statistics, which test the difference between duration levels, were statistically significant,  $Q_B = 13.11$ ,  $p < .005$ ;  $Q_B = 5.86$ ,  $p < .10$ ;  $Q_B = 8.78$ ,  $p < .05$ , for rate, accuracy and comprehension, respectively. These significant  $Q_B$  statistics indicated that there was systematic variability in the treatment effect sizes among the three levels of treatment duration. Specifically, the mean effect sizes for all three outcomes showed that the longer the treatment duration, the smaller the effect sizes.

Table 11. Mean Effect Sizes of Outcomes by Treatment Duration

Outcome	Mean ES	N	95% C. I.		Q <sub>w</sub>	Q <sub>B</sub>
			Lower	Upper		
Reading Rate						13.11 ***
1-6 weeks	0.50	12	0.25	0.75	12.54	
7-20 weeks	0.40	41	0.29	0.51	37.72	
21-36 weeks	0.16	22	0.06	0.26	16.36	
Reading Accuracy						5.86 *
1-6 weeks	0.69	7	0.40	0.98	11.99 *	
7-20 weeks	0.43	16	0.24	0.61	10.36	
21-36 weeks	0.28	9	0.10	0.45	5.17	
Reading Comprehension						8.78 **
1-6 weeks	0.42	9	0.14	0.70	14.47 *	
7-20 weeks	0.43	30	0.31	0.56	25.55	
21-36 weeks	0.18	16	0.05	0.30	18.39	

1. ES: Effect size.

2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.

3. Q<sub>B</sub>: from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.

\* p < .10; \*\*p,.05; \*\*\* p < .005.

### *Treatment Frequency*

Table 12 shows the mean effect sizes and confidence intervals for each outcome at different treatment frequency levels. The mean effect sizes for reading rate and accuracy at all treatment frequency levels were positive, ranging from small to medium effects, and the 95% confidence intervals did not include zero. For reading comprehension, the confidence interval for treatments conducted once to twice a week included zero. The Q<sub>B</sub> statistics for reading comprehension, however, was statistically significant, Q<sub>B</sub> = 15.06, p < .005, indicating that there was systematic variability in the treatment effect sizes among the three levels of treatment frequency. A closer examination indicated that treatments that were conducted

“daily” produced the largest effects (.62) but that there was little difference between the first two levels (i.e., 1-2 times a week, 3-4 times a week) (.26 and .22). For reading accuracy, although with a non-significant  $Q_B$ , treatments conducted daily produced larger effects than treatments conducted less frequently. For reading rate, the pattern seemed to go slightly the other direction, i.e., the more frequent the treatment, the smaller the effects, but one category (1-2 times per week) had only four comparisons.

Table 12. Mean Effect Sizes of Outcomes by Treatment Frequency

Outcome	Mean ES	N	95% C. I.		$Q_w$	$Q_B$
			Lower	Upper		
Reading Rate						
1-2 times a week	0.38	4	0.06	0.71	1.57	0.81
3-4 times a week	0.29	49	0.21	0.38	64.00 *	
Daily	0.24	22	0.10	0.38	13.36	
Reading Accuracy						
1-2 times a week						2.62
3-4 times a week	0.32	18	0.17	0.47	24.01	
Daily	0.51	14	0.34	0.69	6.76	
Reading Comprehension						
1-2 times a week	0.26	4	-0.06	0.58	0.81	15.06 ***
3-4 times a week	0.22	35	0.12	0.32	34.98	
Daily	0.62	16	0.45	0.80	16.34	

1. ES: Effect size.

2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.

3.  $Q_B$ : from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.

\* $p < .10$ , \*\*\* $p < .0005$ .

### *Treatment Format*

Table 13 presents the mean effect sizes and confidence intervals for each outcome with different treatment formats. The mean effect sizes for all three

outcome constructs for all three treatment formats were positive and the 95% confidence intervals did not include zero. For reading comprehension, the  $Q_B$  statistic was statistically significant,  $Q_B=20.00$ ,  $p<.000$ . The significant  $Q_B$  statistic indicated that there was systematic variability in the treatment effect sizes for reading comprehension among the three levels of treatment format. Specifically, teacher-directed instruction produced the largest mean effect (.68 vs. .20 and .38 for tutoring and subject working independently). For reading accuracy, the  $Q_B$  statistic approached significance ( $Q_B=4.53$ ,  $p=.1040$ ), with teacher-directed instruction producing the largest effect (.58 vs. .30 and .35 for tutoring and subject working independently). For reading rate, the  $Q_B$  statistic was non-significant, but the results indicated the opposite direction, with teacher-directed instruction producing the smallest effect (.24 vs. .29 and .46 for tutoring and subject working independently).

Table 13. Mean Effect Sizes of Outcomes by Treatment Format

Outcome	Mean ES	N	95% C. I.		$Q_w$	$Q_B$
			Lower	Upper		
<b>Reading Rate</b>						
Teacher directed instruction	0.24	18	0.09	0.38	10.31	1.32
Tutoring	0.29	51	0.21	0.37	62.41	
Subject independently	0.46	6	0.09	0.84	5.70	
<b>Reading Accuracy</b>						
Teacher directed instruction	0.58	10	0.38	0.78	4.47	4.53
Tutoring	0.30	13	0.14	0.47	14.91	
Subject independently	0.35	9	0.08	0.63	9.49	
<b>Reading Comprehension</b>						
Teacher directed instruction	0.68	13	0.49	0.86	11.34	20.00 ***
Tutoring	0.20	36	0.11	0.30	20.82	
Subject independently	0.38	6	0.06	0.69	6.02	

1. ES: Effect size.

2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.

3.  $Q_B$ : from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.

\*\*\* $p < .000$ .

### *Treatment Focus*

Table 14 presents the mean effect sizes and confidence intervals for each outcome at different treatment focus levels. For reading rate and accuracy, treatments that focused on comprehension only produced the weakest effect, and the 95% confidence intervals included zero. Conversely, for reading comprehension, treatments that focused on fluency only produced the smallest effects and the 95% confidence intervals included zero. For all three outcomes, treatments that focused on fluency and comprehension yielded the largest effect (.35, .56 and .50 for rate, accuracy and comprehension respectively). Further, for comprehension, the  $Q_B$  statistic was statistically significant,  $Q_B = 16.27$ ,  $p < .0005$ , and that for accuracy was approaching significance,  $Q_B = 4.47$ ,  $p = .1069$ , suggesting that systematic variability existed in the effect sizes among the three levels of treatment focus for these outcomes. Specifically, interventions that focused on both fluency and comprehension produced the largest mean effect size (.56 for accuracy and .50 for comprehension).

Table 14. Mean Effect Sizes of Outcomes by Treatment Focus

Outcome	Mean ES	N	95% C. I.		Q <sub>w</sub>	Q <sub>B</sub>
			Lower	Upper		
Reading Rate						3.08
Fluency	0.26	42	0.17	0.36	48.96	
Comprehension	0.12	7	-0.12	0.37	4.13	
Fluency+Comprehension	0.35	26	0.23	0.46	23.56	
Reading Accuracy						4.47
Fluency	0.34	17	0.18	0.49	22.33	
Comprehension	0.16	3	-0.27	0.59	0.32	
Fluency+Comprehension	0.56	12	0.36	0.75	6.27	
Reading Comprehension						16.27 ***
Fluency	0.13	22	0.00	0.26	16.87	
Comprehension	0.26	7	0.03	0.48	8.52	
Fluency+Comprehension	0.50	26	0.37	0.62	25.53	

1. ES: Effect size.

2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.

3. Q<sub>B</sub>: from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.

\*\*\*p < .0005.

### *Treatment Session Length*

Table 15 presents the mean effect sizes and confidence intervals for each outcome at different treatment session length levels. Again, the mean effect sizes for all three outcome constructs for all treatment session length levels were positive and the 95% confidence intervals did not include zero. And for reading comprehension, the Q<sub>B</sub> statistic was statistically significant, Q<sub>B</sub> = 11.91, p < .005, indicating that systematic variability exists in the effect sizes for reading comprehension among the three levels of treatment session length. Interventions that lasted more than 45 minutes produced the largest mean effect size (.70). However, caution is needed because there were only 4 effect sizes in this category. For reading rate and

accuracy, the  $Q_B$  statistics were non-significant. For all three outcome measures, interventions that lasted between 31-45 minutes seemed to have produced the smallest mean effect sizes.

Table 15. Mean Effect Sizes of Outcomes by Treatment Session Length

Outcome	Mean ES	N	95% C. I.		$Q_w$	$Q_B$
			Lower	Upper		
Reading Rate						1.83
15-30 minutes	0.33	30	0.21	0.45	33.00	
31-45 minutes	0.24	41	0.15	0.33	44.59	
> 45 minutes	0.44	4	0.11	0.77	1.38	
Reading Accuracy						1.87
15-30 minutes	0.47	23	0.31	0.63	25.82	
31-45 minutes	0.28	6	0.11	0.46	7.90	
> 45 minutes	0.43	3	0.06	0.80	1.18	
Reading Comprehension						11.91 ***
15-30 minutes	0.44	26	0.30	0.59	26.40	
31-45 minutes	0.20	25	0.09	0.30	28.68	
> 45 minutes	0.70	4	0.36	1.04	3.58	

1. ES: Effect size.

2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.

3.  $Q_B$ : from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.

\*\*\*p < .005.

### *Treatment Intensity*

Table 16 presents the mean effect sizes and confidence intervals for each outcome at different treatment intensity levels. The mean effect sizes for both reading rate and accuracy are all positive and the 95% confidence intervals do not include zero. For reading comprehension, while all the mean effect sizes are positive, one of the 95% confidence intervals included zero. Further, the  $Q_B$  statistics

for accuracy ( $Q_B = 5.38, p < .10$ ) and comprehension ( $Q_B = 19.30, p < .0001$ ) were significant (see Table 16).

Table 16. Mean Effect Sizes of Outcomes by Treatment Intensity

Outcome	Mean ES	N	95% C. I.		$Q_w$	$Q_B$
			Lower	Upper		
Reading Rate						1.57
0-60 minutes/week	0.40	11	0.18	0.61	14.71	
61-120 minutes/week	0.26	42	0.17	0.35	41.73	
> 120 minutes/week	0.31	22	0.17	0.45	21.72	
Reading Accuracy						5.38 *
0-60 minutes/week	0.59	7	0.24	0.94	12.78 **	
61-120 minutes/week	0.28	15	0.13	0.44	11.15	
> 120 minutes/week	0.55	10	0.34	0.75	4.08	
Reading Comprehension						19.30 ***
0-60 minutes/week	0.18	7	-0.11	0.47	3.01	
61-120 minutes/week	0.22	35	0.12	0.32	32.97	
> 120 minutes/week	0.68	13	0.50	0.87	11.91	

1. ES: Effect size.

2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.

3.  $Q_B$ : from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.

\*  $p < .10$ ; \*\*  $p < .05$ ; \*\*\*  $p < .0001$ .

### *Total Treatment Time*

Table 17 presents the mean effect sizes and confidence intervals for each outcome at different levels of total treatment time. For reading rate and comprehension, the mean effect sizes at different treatment time levels were all positive and the 95% confidence intervals did not include zero. Furthermore, the  $Q_B$  statistic for comprehension was statistically significant,  $Q_B = 10.89, p < .05$ . For reading accuracy, the mean effect sizes were all positive, but the 95% confidence intervals included zero and the  $Q_B$  statistic was non-significant (see Table 17).



Table 17. Mean Effect Sizes of Outcomes by Total Treatment Time

Outcome	Mean ES	N	95% C. I.		Q <sub>w</sub>	Q <sub>B</sub>
			Lower	Upper		
Reading Rate						10.89 **
1-15 hours	0.34	18	0.15	0.53	17.75	
15.01-30 hours	0.49	17	0.34	0.65	16.56	
30.01-40 hours	0.19	33	0.09	0.29	24.48	
> 40 hours	0.24	7	0.04	0.43	10.06	
Reading Accuracy						3.27
1-15 hours	0.49	13	0.27	0.70	20.20 *	
15.01-30 hours	0.83	2	0.10	1.57	0.43	
30.01-40 hours	0.37	14	0.23	0.52	9.17	
> 40 hours	0.16	3	-0.27	0.59	0.32	
Reading Comprehension						1.02
1-15 hours	0.31	14	0.09	0.53	18.07	
15.01-30 hours	0.34	15	0.18	0.50	9.94	
30.01-40 hours	0.28	21	0.17	0.40	35.63 **	
> 40 hours	0.45	5	0.12	0.77	2.53	

1. ES: Effect size.

2. Mean effect sizes and confidence intervals were from results from a fix effects model analysis.

3. Q<sub>B</sub>: from a fixed effects model to assess ES heterogeneity beyond subject level sampling error.

\*p < .10, \*\*p < .05.

## Multiple Regression on the Variability in the Effect Sizes

### *Predictor Variables*

Conceptually, three categories of variables were considered potentially meaningful predictors of the variability in the distribution of the effect sizes of the three outcomes: variables representing subject characteristics, those representing the type of treatments, and those representing treatment characteristics or features.

### *Subject Characteristics Predictors*

The category of subject characteristics contained the following six variables,

age, grade level, reading level, race, sex, and student type. Student type indicated whether the sample was composed of students with disabilities, remedial readers, normal readers, or mixed readers (i.e., sample had readers with disabilities, remedial readers and normal readers). Student type was coded as following: 1=students with disabilities, 2=remedial readers; 3=mixed, with students with disabilities, remedial and normal readers, 4=normal readers. Reading level indicated the reading ability of the subjects at the beginning of treatment, whether at the current grade level, one year below, two years below, or three years below current grade level. Ten of the 84 (or 12%) cases did not report the reading level of the subjects and, as a result, reading level had to be excluded from analyses. Further, since almost half of the studies did not report information regarding subjects' sex, ethnicity, or age, examination of the relationships between these characteristics and the effect sizes was not attempted.

#### *Treatment Characteristics Predictors*

The treatment characteristics category contained seven variables: treatment duration (1=1-6 weeks, 2=7-20 weeks, 3=21-36 weeks), treatment frequency (1=1-2 times per week, 2=3-4 times per week, 3=daily), treatment format (1= teacher-directed whole class or small group instruction, 2= tutoring - adults, peers or cross age, or 3= students working independently), length of each treatment session (1=15-30 minutes, 2=31-45 minutes, 3=46-60 minutes), treatment intensity (1=0-60 minutes per week, 2=61-120 minutes per week, 3=more than 120 minutes per week), total treatment time (1=0-15 hours, 2=16-30 hours, 3=31-40 hours, 4=more than 40 hours) and treatment focus (fluency only, comprehension only or both

fluency and comprehension). Treatment focus was dummy-coded as following to compare fluency only and comprehension only against fluency + comprehension: treatment focus is fluency only (1= fluency, 0= comprehension, 0= fluency + comprehension), or treatment is comprehension only (1= comprehension, 0=fluency, 0=fluency + comprehension).

### *Treatment Type Predictors*

As indicated earlier, treatment type was coded in three different ways to reflect 1) specific fluency building strategies, 2) different text levels, and 3) different practice levels.

First, it was coded to reflect specific fluency building strategies. These strategies included repeated reading or other guided oral reading of connected texts, repeated reading or other guided oral reading of mixed texts, encouraging children to read more, general teaching strategies or programs, and word level training. For the multiple regression analyses, this variable was converted into two codes of fluency strategy: fluency strategy code A and fluency strategy code B. In fluency strategy code A, the first two (of the five) categories were collapsed as “repeated reading and other guided oral reading, while the last three (of the five) categories were collapsed as “other methods” to get 1=repeated reading and other guided oral reading, and 2=other methods.

Fluency strategy code B was created so it was also possible to examine whether there existed systematic variability between either of the repeated reading/guided oral reading strategies and other methods. The codes followed: 1=repeated reading/other guided oral reading of connected texts, 2= repeated

reading/other guided oral reading of mixed texts, 3=other methods. This variable was dummy coded to contrast the first two categories against “other methods” as following: Strategy is repeated reading/other guided oral reading of connected texts (1=repeated reading/guided oral reading of connected texts, 0=repeated reading/guided oral reading of mixed texts, 0=other methods); strategy is repeated reading/other guided oral reading of mixed texts (1=repeated reading/guided oral reading of mixed texts, 0=repeated reading/guided oral reading of connected text, 0=other methods).

Second, the interventions were coded in terms of different text levels of the treatment content, including non-connected text level (sounds, words and phrases), connected text level (stories), and a mixed level, i.e., level of both connected texts and non-connected texts (sounds, words and stories combined). This code (text level of treatment) was scaled as: 1=treatment is at non-text level; 2= treatment is at both non-text and text level; 3= treatment is at text level.

Third, the interventions were coded in terms of how they were conducted, including non-repetitive practice, repetitive practice of non-connected texts (sounds, words and phrases), repetitive practice of connected texts (stories), repetitive practice of both connected and non-connected texts (sounds, words, phrases, and stories). This was further collapsed to include only two levels, non-repetitive practice and repetitive practice. For clarity purposes, this coding of treatment was labeled as following:

Repetitiveness/text of treatment: 1= treatment is non-repetitive; 2= treatment is repetitive at non-text level; 3=treatment is repetitive at both text and non-text level;

4= treatment is repetitive at text level.

Repetitiveness of treatment: 1= treatment is non-repetitive; 2=treatment is repetitive.

Altogether fifteen variables could be entered as potential predictors of the effects on reading fluency, as listed below:

Subject characteristics:

- Subject characteristic/student type

Treatment type:

- Fluency strategy type A (repeated reading/oral reading vs. other methods)
- Fluency strategy type B: RR/OR of stories (repeated reading/oral reading of stories vs. other methods)
- Fluency strategy B: RR/OR of mixed texts (repeated reading/oral reading vs. other methods)
- Repetitiveness of treatment
- Text level of treatment
- Repetitiveness/text level of treatment

Treatment characteristics:

- Treatment duration
- Treatment focus: fluency (fluency vs. fluency + comprehension)
- Treatment focus: comprehension (comprehension vs. fluency + comprehension)
- Treatment session length
- Treatment format

- Treatment frequency
- Treatment intensity
- Total treatment time

To guide the analyses so that meaningful predictors were entered in the multiple regression analyses, zero-order correlations between predictors and effect sizes as well as inter-correlations among predictors were computed (see Tables 18 and 19 for correlations). Some predictors were highly correlated with each other, while others were not correlated with the effect sizes of the three outcome variables. Therefore, some of the variables were excluded from the multiple regression analyses for the three outcomes.

Table 18. Correlations between Effect Sizes and Predictor Variables

Variables	ES for Rate	Variables	ES for Accuracy	Variables	ES for Comp.
Fluency Str. A	-.38***	TX repetitiveness	.52***	TX intensity	.43***
RR/OR-mixed texts	.28*	TX repetitive & text	.44*	TX repetitive & text	.37**
TX repetitiveness	.25*	Subj. character.	-.40*	TX repetitiveness	.31*
TX repetitive & text	.23*	TX text level	-.30	TX focus-Fluency	-.30*
TX frequency	-.22	TX duration	-.28	TX format	-.29*
TX format	.21	RR/OR-mixed text	.25	Fluency Str. A	-.28*
TX focus-comp.	-.21	Fluency Str. A	-.21	TX frequency	.25
TX duration	-.19	TX focus-Comp	-.20	TX total time	.24
TX intensity	-.14	TX format	-.13	RR/OR-mixed text	.12
Subj. character.	-.12	TX total time	-.13	RR/OR-stories	.10
TX total time	-.10	RR/OR-stories	-.10	TX text level	.07
TX focus-fluency	.09	TX frequency	.06	TX session length	.05
RR/OR-stories	.05	TX focus-fluency	-.03	Subj. character.	-.05
TX session length	-.05	TX session length	-.03	TX duration	-.03
TX text level	.00	TX intensity	.00	TX focus-Comp	.01

Comp.: Comprehension. \*p<.05; \*\*p<.01; \*\*\*p<.005

Table 19. Correlations between Predictor Variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Subj. Character.														
2. TX Session Length	-.10													
3. TX Duration	.10	.36												
4. TX Frequency	.02	-.05	-.17											
5. Fluency Str. A	-.05	.44**	.42**	-.18										
6. RR/OR of stories	.17	-.32**	-.33**	.09	-.39**									
7. RR/OR of mixed texts	.08	.47**	.41**	-.23*	-.51**	.59**								
8. TX repetitiveness	.17	.13	.06	-.23*	.32**	-.09	.46**							
9. TX text level	.15	-.12	-.10	.19	-.64**	.77**	-.43**	-.36**						
10. TX repetitive & text	.15	.08	.00	-.05	-.03	.31**	.24*	.86**	.13					
11. TX Format	.11	.27**	.02	-.29**	-.01	.03	.01	-.13	.02	-.19				
12. TX Focus-F <sup>a</sup>	-.19	-.18	.01	-.27*	-.11	-.03	.13	.28*	-.20	.14	.42**			
13. TX Focus-C <sup>b</sup>	-.02	.15	.23*	.18	.33**	-.23*	-.08	.26*	.02	-.38**	.09	-.36*		
14. TX Intensity	.06	.41**	.07	.59**	.14	-.15	.10	-.11	.10	-.10	-.51**	-.38**	.17	
15. TX Total Time	.03	.53**	.67**	.05	.35**	-.32**	.31*	-.08	.01	-.05	-.22*	-.26**	-.11	.62**

<sup>a</sup>Treatment focus is fluency only.

<sup>b</sup>Treatment focus is comprehension.

Flu. Str. A: 1=Repeated reading and other guided oral reading; 2=Other methods

\*p<.05; \*\*p<.01.



For reading rate, fourteen variables were included, namely, fluency strategy type A, repeated reading/oral reading of mixed texts, repetitiveness of treatment, repetitiveness/text level of treatment, treatment frequency, format, treatment focus-comprehension, treatment duration, treatment intensity, student type, total treatment time, treatment focus-fluency, repeated reading /oral reading of connected texts, treatment session length.

For reading accuracy, fourteen variables were included: Repetitiveness of treatment, repetitiveness/text level of treatment, student type, text level of treatment, treatment duration, repeated reading/guided oral reading of mixed texts, fluency strategy type A, treatment focus-comprehension, treatment format, total treatment time, repeated reading/guided oral reading of connected texts, treatment frequency, treatment focus-fluency, and treatment session length.

For comprehension, all fifteen variables were included. However, repetitiveness of treatment and repetitiveness/text of treatment, treatment duration and total treatment time, total treatment time and treatment intensity highly correlated with each other, so they were not entered in the analyses at the same time.

To analyze the variability of effect sizes as a function of study characteristics, a weighted, fixed effects multiple regression was used with the effect sizes for reading rate, accuracy, and comprehension separately with appropriate predictor variables described above. With all three kinds of effect size estimates (i.e., reading rate, reading accuracy, and reading comprehension), the following general procedures were used in the multiple regression analysis. First, all potential predictors were entered in an analysis. Second, if the results from the analysis did not appear to be the best fit, the weakest predictor was dropped from the analysis, and the analysis was repeated until the best possible model was obtained. The final results are presented below.

### *Multiple Regression Results for Effect Sizes of Reading Rate*

For effect size estimates of reading rate, treatment duration, repeated reading/guided oral reading of mixed texts, student type, repeated reading/guided oral reading of connected texts, and repetitiveness of practice in treatment (repetitive vs. non-repetitive) were significant predictors, indicating that the effectiveness of the fluency promoting interventions on reading rate differed as a function of these factors. The model accounted for a significant proportion of the observed variability in the effect size estimates ( $R^2 = .41$ ,  $p < .000$ ) and left a non-significant residual (see Table 20). Treatment effect differed as fluency strategies used differed. Specifically, repeated reading and other guided oral reading, regardless of the type of text involved, produced larger effect than other types of fluency strategies or programs. Treatment effects also appeared to differ as a function of the type of treatment received; specifically, repetitive practice of tasks produced larger effects than non-repetitive practice of tasks (.30 vs. .15 for repetitive and non-repetitive). However, treatment duration was negatively associated with treatment effects, suggesting that the longer the treatment, the smaller the effects (see Table 20), which was evidenced by the mean effect sizes at different duration levels (.50 vs. .40 vs. .16 for 1-6 weeks, 7-21 weeks and 21-36 weeks, respectively, see Table 11). Further, treatment effects appeared to be dependent on the type of students who were involved in the study; treatments seemed to have yielded greater effects on reading rate for students with disabilities (.34) and remedial readers (.37), and have produced the smallest effects for normal readers (.18) (see Tables 20 & 7).

### *Multiple Regression Results on Effect Sizes of Reading Accuracy*

For effect size estimates of reading accuracy, the multiple regression results suggest four significant predictors: repetitiveness of practice (i.e., repetitive practice vs. non-

repetitive practice), treatment duration, student type, and treatment focus-comprehension only. The model accounted for a significant proportion of the observed variability in the effect sizes ( $R^2 = .44$ ,  $p < .01$ ) and left a non-significant residual (see Table 20). Results suggested that students with disabilities had the largest mean effect size (.86), while normal readers had the smallest mean effect size (.26). Results also suggested that (a) the longer the treatment, the smaller the effect, and (b) treatment was more effective if it was repetitive in nature, be it repetitive practice of sounds, words or stories, or a combination of these. It is therefore not surprising that students who repeatedly practiced reading sounds, words/phrases, stories, or all text levels tended to read with greater accuracy than those who engaged in non-repetitive reading. In addition, results suggested that if the treatment focused on comprehension only, there was a slight negative effect on reading accuracy.

#### *Multiple Regression Results on Effect Sizes of Reading Comprehension*

For reading comprehension, six variables, i.e., treatment intensity, repetitiveness of treatment, student type, treatment focus-fluency, text level of treatment, and treatment session length were found to be significant predictors. The model accounts for 49% of the variability in the effect size estimates ( $R^2 = .49$ ,  $p < .000$ ). The residual was non-significant (see Table 20). Results indicated that the more intense the treatment (as indicated in minutes spent on treatment each week), the greater the effects, and that treatments with an emphasis on both fluency and comprehension yielded significantly better effects than treatments with an emphasis on only fluency, and that repetitive practice of tasks produced larger effects on reading comprehension than non-repetitive practice of tasks. Further, results showed that interventions involving stories produced larger effects on comprehension than those involving non-story materials. Lastly, results showed that

students with disabilities or remedial readers enjoyed larger effects (.32 and .37 respectively) than normal readers (.20) (see Table 20 & Table 7).

Table 20. Final Results of Weighted Multiple Regression Analyses

Predictors for	B	95% C. I.		Beta
		Lower	Upper	
<b>Reading Rate<sup>a</sup></b>				
Treatment duration	-0.26 ***	-0.39	-0.13	-0.51
RR/OR of mixed texts	0.38 ***	0.15	0.57	0.55
Student type	-0.11 ***	-0.18	-0.04	-0.38
RR/OR of connected texts	0.21 *	-0.03	0.44	0.27
Repetitiveness of treatment	0.21	-0.09	0.51	0.17
<b>Reading Accuracy<sup>b</sup></b>				
Repetitiveness of treatment	0.48 **	0.08	0.87	0.53
Treatment duration	-0.20 **	-0.37	-0.03	-0.43
Student type	-0.12 **	-0.23	-0.01	-0.38
Treatment focus-comp.	-0.44	-0.17	-0.05	-0.33
<b>Reading Comprehension<sup>c</sup></b>				
Treatment intensity	0.31 ***	0.14	0.49	0.47
Repetitiveness of treatment	0.31 **	0.04	0.57	0.31
Student type	-0.09 **	-0.17	-0.01	-0.26
Treatment focus-fluency	-0.16 *	-0.35	0.02	-0.23
Text level of treatment	0.15	-0.05	0.34	0.22
Treatment session length	-0.12	-0.29	0.06	-0.19

\*  $p < .10$ ; \*\*  $p < .05$ ; \*\*\*  $p < .01$ ;

<sup>a</sup> Model:  $Q(5) = 32.74$ ,  $p < .0000$ ; Residual:  $Q(69) = 47.00$ , *ns.*  $R^2 = .41$

<sup>b</sup> Model:  $Q(4) = 14.82$ ,  $p < .01$ ; Residual:  $Q(27) = 18.57$ , *ns.*  $R^2 = .44$

<sup>c</sup> Model:  $Q(6) = 32.61$ ,  $p < .000$ ; Residual:  $Q(48) = 34.57$ , *ns.*  $R^2 = .49$

## CHAPTER V

### DISCUSSION AND CONCLUSIONS

#### Limitations

This meta-analysis has several limitations. First, the number of studies included was small and as a result, might have resulted in low statistical power, which in turn might have led to the non-significant  $Q_B$  statistics for reading rate, accuracy and comprehension. The non-significant  $Q_B$  statistics compromised the justification for all subsequent analyses. Further, the majority of this small set of studies did not provide information regarding subjects' sex, ethnicity, age, and reading level, precluding the possibility to examine the relationships between treatment effects and characteristics of subject as represented by these variables.

Second, within this set of studies, many different instructional procedures were examined, so many that it is impossible to determine the best procedures based on the few studies that examined the procedures. No method was used so often that a reliable estimate of effect size would be possible. Also, variations across studies are subtle in terms of material selection and amount or type of treatment. Some treatments were delivered by teachers or researchers, some by other adults, some by other students, and some by students themselves with computers or tape recorders. This wide variability in treatments created difficulty in categorizing them and finding a common thread to logically and conceptually analyze the effect sizes. Even though I categorized specific fluency building strategies into several groups

such as a) repeated reading and other guided oral reading of connected text, b) repeated reading and other guided oral reading of mixed text, c) encouraging students to read more, d) word recognition or decoding training, and e) other general instruction methods and programs, the distinction was far from clear cut. In addition, there were very few studies in the last three categories, and there was considerable variability in the category of repeated reading and other guided oral reading of mixed text, as evidenced by the significant within study Q statistic ( $Q_w=53.07$ ,  $p<.05$ ). For lack of a better way to categorize them, I put these strategies in one category based on the practice they provided and text involved. Therefore, while comparisons and contrasts were made among these fluency building strategies, it seems that there were consistent differences among them for the three outcome constructs.

Third, studies that met the inclusion criterion of involving a control group had to be eliminated from the meta-analysis because of insufficient data. As a result, the studies included may be susceptible to the criticism and question as to how representative they were of the existing body of literature on fluency research for elementary students.

Fourth, this meta-analysis does not examine the role of prosody in the development of reading fluency and the ways in which it contributes to the development of reading comprehension. A meta-analysis requires studies to quantify and measure student outcomes, however, prosody is not easily quantifiable and measurable. It consists of a series of features including pitch or intonation, stress or emphasis, and tempo or rate and the rhythmic patterns of language, all of which contribute to an expressive rendering of a text (Allington, 1983a; Dowhower, 1991).

In addition, prosody includes appropriately chunking groups of words into phrases or meaningful units in accordance with the syntactic structure of the text. All these features are more difficult to quantify than reading rate, accuracy and comprehension. This probably explains why prosody was not examined in most studies included in the present meta-analysis.

Fifth, given that the ultimate goal of reading is the construction of meaning (Anderson, Hiebert, Wilkinson, & Scott, 1985), it is important to assess the role fluency plays in comprehension. Does fluency (rate and accuracy) influence comprehension? However, due to the small number of studies, and that many studies only measured fluency without measuring comprehension, no attempt was made to examine whether the development of fluency skills affected the development of reading comprehension.

Sixth, many studies in the current meta-analysis did not report the fidelity of treatment implementation; it was therefore difficult to determine how accurately the various interventions in the studies were implemented and how much of the effectiveness (or ineffectiveness) of the interventions could be explained by program or treatment implementation.

Despite these limitations, the present meta-analysis has helped answer the research questions it set out to answer and has yielded interesting and possibly practical, statistically significant findings. The information coded for the studies has helped answer whether fluency building interventions improve reading fluency and comprehension; whether the average effect sizes of different interventions were significantly different from zero; whether those effect sizes differed from each other;

and what variables predicted the effect of various interventions on reading rate, accuracy and comprehension. The analyses also shed light on designing and developing fluency building programs and strategies, provided practical guides to teaching fluency, and identified interesting angles for further exploration and questions for future research.

## Summary and Discussion of Findings

### *Overall Effects of Fluency Building Strategies*

Mean effect sizes of various treatments for reading rate, accuracy, and comprehension were all within the range of medium effects (.30 -.41) and were all significantly different from zero. This seems to suggest that the various interventions are effective and robust in general. In other words, treatment effectiveness does not seem related to a particular type of treatment. This finding is consistent with the finding of NRP (2000) report regarding the effectiveness of repeated reading and other guided repeated oral reading. The authors noticed the robustness of various fluency building procedures and a lack of clear differences among these procedures. However, upon further examination, there are differences in treatment effects by specific fluency strategies. Compared to fluency building strategies such as word recognition or decoding training alone, or teaching methods labeled as effective teaching and direct instruction in general, the category of repeated reading and other guided oral reading seemed the most effective, and among repeated reading and other guided oral reading practices, the ones involving connected texts seemed



more effective than those involving both words and connected texts. The strategies in this category include repeated reading (Samuels, 1979), neurological impress (Henckelman, 1969), paired reading (Topping, 1989), shared reading, assisted reading, to name a few. The purpose of these procedures is to help students through oral reading practice and guidance to develop fluent reading habits that would allow them to read text more quickly, accurately, and with appropriate expression and understanding. These procedures share the following distinct features: a) students read a text over and over for a certain number of times or until a pre-specified level of proficiency has been reached; b) increase the amount of oral reading practice through the use of one-to-one instruction, tutors, audiotapes, peer guidance, or other means; and c) some have carefully designed feedback routines to monitor the reader's performance.

In addition, for all three outcomes, treatments with built-in repetitive practice yielded greater effects than treatments without built-in repetitive practice (.15 vs. .30, .14 vs. .46 and .20 vs. .34 for rate, accuracy and comprehension respectively, see Table 9.2). This finding is consistent with the previous finding. Many of the treatments with built-in repetitive practice fall within the category of repeated reading and guided oral reading. Also, for all three outcomes, interventions involving remedial readers and students with disabilities yielded larger effects than interventions involving normal readers.

### *Factors Affecting Effects of Treatments*

Results from the weighted multiple regression analyses indicated that there were two common variables that predicted the variability in the effect size distributions of reading rate, accuracy, and comprehension: repetitiveness of treatment and student type. For all three outcome constructs, treatments that provided repetitive practice for students produced larger effects than those that did not provide repetitive practice. When students repeatedly read their material, whether it was sounds, words, stories, or a combination of all, they tended to read faster and with greater accuracy and comprehension than students who practiced reading their stories without repetition. Through repeated and extended practice in which large quantities of material are read, a reader develops fluency skills that go beyond accuracy of recognition to automaticity of recognition (Allington, 1977, 1984; Snow, Burns, & Griffin, 1998), and the reader becomes so fluent in reading that words are recognized accurately, quickly, and with ease and that the text sounds like spoken language when read aloud.

Student type is the second common predictor for all three outcome constructs. Treatments involving remedial students and students with disabilities yielded greater effects than treatments involving normal readers. This finding may be due to the possibility that researchers have used easier materials and outcome measures with remedial readers and students with disabilities, resulting in artificially inflated treatment effects.

In addition to the two common predictors, each outcome construct also has its own unique predictors. For reading rate, the unique predictors are treatment

duration, repeated reading or other guided oral reading of mixed texts, and repeated reading and other guided oral reading of stories. However, for treatment duration, the prediction is in a direction that does not seem to make much sense, i.e., the longer the treatment lasted, the worse the students' reading rate. Compared to general fluency building methods such as effective teaching and direct instruction, the category of repeated reading and guided repeated oral reading of stories or of mixed texts (i.e., words, phrases and stories) produced larger effects.

For reading accuracy, the unique predictors were treatment duration and treatment focus-comprehension only. Again, results showed that the longer the treatment, the smaller the effects. Although it is natural to assume that longer programs produce larger effects on achievement, this is not always the case. Some researchers have found that programs shorter than 44 hours and longer than 210 hours have resulted in very small negative effects, and that the ideal program length for improving reading achievement appears to be between 44 and 84 hours (Miller & Snow, 2004). While these findings do not agree exactly with what the present meta-analysis has revealed, it is not surprising, then, that this meta-analysis discovered that interventions with treatment duration equal to or shorter than 20 weeks produced larger effects than those that lasted longer than 20 weeks. In addition, implementation slippage is more likely to occur when treatments are of considerable duration.

If the finding regarding treatment duration seems a little puzzling, the finding that, compared to treatments focusing on both fluency and comprehension, treatments focusing on comprehension only had a negative effect on reading

accuracy is not only conceivable but expected. As Faulkner and Levy (1999) have speculated, there could be a hierarchical order to reading development, with the lowest level readers improving in word recognition and the highest in comprehension. Too much emphasis on higher level skills training (such as comprehension) when readers are still developing lower level skills (such as word recognition or fluency) may adversely affect the development of lower level skills. Or, according to the resource-based fluency theory, the opposite could be true as well: A reader has limited mental capacity or resources, and when too much attention is spent on resource-demanding comprehension activities, there is little capacity left for developing fluency. All these possibilities may suggest that, before a reader achieves fluency, too much emphasis on reading comprehension training may actually hinder a reader's fluency development.

On the other hand, a word of caution can be drawn from a short-term study (Anderson, Wilkinson, & Mason, 1991, as cited in the NRP report, 2000) that found that too much attention to fluency issues within a reading lesson could detract from reading comprehension. It should be noted that in all the studies in this meta-analysis, the fluency intervention was only part of the instruction that students received. In most cases, the fluency intervention was relatively brief (15 to 30 minutes per lesson), and students who received these lessons were still engaged in other reading activities including comprehension instruction. Guided repeated oral reading and repeated reading provide students with practice that substantially improves fluency and, to a lesser extent, reading comprehension. They appear to do

so, however, in the context of an overall reading program, not as stand alone interventions.

For reading comprehension, the unique predictors were treatment intensity, treatment focus-fluency only, text level of treatment, and treatment session length. Treatments with greater intensity (i.e., with a greater amount of time spent on treatment per week) produced larger effects than those without. And treatment involving connected texts also yielded larger effects than those involving non-connected or mixed texts. Further, results showed that if the treatment focused on fluency only (as compared to treatments that focused on both fluency and comprehension), or if the length of each treatment session was longer, the treatment effect was weaker. This suggests that to improve reading comprehension, it is important to include both a fluency and comprehension component in the treatment.

To conclude, when compared to the traditional instruction (i.e., the business as usual type of instruction in the control conditions), the results seem clear: The various interventions included in this meta-analysis are moderately effective in improving reading fluency and comprehension at the elementary level, regardless of grade level. Further, these meta-analytic results indicate that the best intervention may be one in which a reading comprehension training component is included. Since the ultimate goal of reading is comprehension, this finding is especially important for researchers designing and developing fluency building programs or interventions. When different approaches to fluency instruction are compared, however, the results are less clear. It seems, nonetheless, that repeated reading and other guided oral reading, compared to general teaching strategies as indicated by

methods such as direct instruction or effective teaching, are more effective in helping students develop fluency and comprehension.

## Implications for Practice, Fluency Theories, and Research

### *Implication for Reading Practice*

Increasingly, teacher educators, educational researchers, and theorists have called for more attention to direct instruction in fluency. Various procedures have been proposed for teaching students to read quickly, accurately, and with proper expression, though it is evident that this remains a serious weakness among many school children.

A number of different approaches to developing reading fluency were covered by the studies in this meta-analysis, including repeated readings, assisted reading, the Oral Recitation Lesson (Hoffman, 1987), etc. These approaches were varied and no method was used so often that a reliable estimate of effect size would be possible. However, a few interesting and practically significant findings did emerge.

### *Repetitive Practice Vs. Non-Repetitive Practice*

As discussed earlier, automaticity theorists suggest that it is through extended and repeated practice that readers develop reading fluency. The studies in the current meta-analysis provide a persuasive case attesting to the soundness of the belief: Interventions in which students read passages orally multiple times while receiving guidance or feedback from peers, parents, or teachers are effective in improving reading fluency (rate and accuracy) and comprehension. These

procedures are not particularly difficult to use; nor do they require a lot of special equipment or materials, although it is uncertain how widely used they are at this time. These procedures help improve students' reading ability, and they help improve the reading of remedial students and students with disabilities or learning problems. Therefore, it is clear that when developing fluency interventions or programs, researchers should include repetitive practice as one of the most important features. When teaching reading, teachers need to provide students with ample practice in oral reading, through peer tutoring, corrective feedback, repetitive practice to enhance fluency and comprehension.

#### *Focus On Fluency and Comprehension*

Comprehension is critically important to the development of children's reading skills and therefore their ability to obtain an education. Indeed, reading comprehension has come to be viewed as the "essence of reading" (Durkin, 1993), essential not only to academic learning but to life-long learning. Although it is not clear from the studies in the current meta-analysis what role the development of fluency plays in the development of reading comprehension, the studies provide a persuasive case underscoring the importance of including a comprehension component in treatments. They demonstrate that fluency building interventions with an added component of comprehension training are more effective in improving both student reading fluency and especially, reading comprehension than those focusing on fluency only or comprehension only. It appears that oral reading practice and feedback or guidance is most likely to influence measures that assess reading rate and reading accuracy. Nevertheless, the impact of these procedures on

comprehension is significant as well. These changes in comprehension might take place simultaneously, with the improvements in fluency mediating the improvements in comprehension, or there could be a hierarchical order to this, as Faulkner and Levy (1999) have speculated, with the lowest level readers improving in word recognition and the highest in comprehension. This points out that when researchers and practitioners design, develop reading programs, they should add a reading comprehension component in their programs. It also points out to teachers that when teaching reading, they should emphasize both fluency and comprehension.

#### *Differential Treatment Effects For Students With Disabilities*

The studies in the current meta-analysis show that interventions involving remedial readers and students with disabilities produced larger effects than interventions involving normal readers. This finding, however, does not necessarily mean that interventions involving remedial readers and students with disabilities are more effective than those involving normal reader. Instead, we should examine other factors associated with the studies, such as the reading level of the subjects at the time of treatment and training materials used in the treatment. Typically, researchers choose to use materials that are at students' reading level, and conceivably for remedial readers and students with disabilities, the training materials are easier than those used with normal readers and subsequently the measures or tests may be easier than those for normal readers. In other words, students with disabilities and remedial readers might be reading easier materials, resulting in artificially elevated effects. This highlights the importance of documenting the characteristics of training materials, such as level of difficulty and type of material, and the importance of



including standardized outcome measures. The researcher might be able to get a better picture of program effectiveness (i.e., how generalizable or reliable the results are) by using more standard and reliable measures.

### *Implications for Fluency Theories*

As discussed in Chapter II, there are three major theories about fluency development: Limited mental resources theory by LaBerge and Samuels (1974), the interactive-compensatory model theory by Stanovich (1980), and instance theory or information encapsulation by Logan (1997).

According to LaBerge and Samuels, an individual has a limited amount of attention available for any given cognitive task. Therefore, attention used for one activity means attention unavailable for another. In the case of reading, an individual has to perform at least two tasks: determining what words constitute the text while simultaneously constructing meaning. The greater the amount of attention expended on decoding the words, the less attention there is available for comprehension. In order for a reader to have enough attention for comprehension, it is necessary for a reader to develop decoding to a point where each word is recognized instantaneously. Therefore ideally, readers should recognize words automatically. If they do not, they must rely on contextual information. As more mental resources are used for contextual analysis in order to identify words, fewer of these resources are available for constructing meaning. It is only when readers achieve automaticity in word recognition that they can concentrate on the meaning of text.

There are two parts in the interactive-compensatory model (Stanovich, 1980): The interactive model and the compensatory model. According to the interactive model, information from multiple sources (e.g., feature extraction, orthographic, phonological, semantic, and syntactic) is available to aid readers in their comprehension of text. In this model, higher level processes constrain the alternatives at lower levels but are themselves constrained by lower level processes. According to the compensatory model, deficiencies in processes at a particular level in the processing hierarchy can be compensated by a greater use of information from other levels, regardless of level of deficiencies. Therefore, according to interactive-compensatory model, higher-level processes can actually compensate for deficiencies in lower-level processes; bottom-up (text-driven or word processing) and “top-down” (meaning-driven or hypothesis forming) processes operate simultaneously.

Instance theory or information encapsulation (Logan, 1997) caused a shift in fluency theory from emphasis on cognitive resource use to emphasis on knowledge representation. It stresses the development of a knowledge base, through which the reader codes information without considering alternative possibilities about what the word means. The key causal property is the development of a high-quality representation of knowledge in memory that allows automatic access.

Despite these theoretical differences about fluency development, the question remains the same for every researcher, “How do readers become fluent or automatic?” All three automaticity theories suggest that it is through practice and repetition in consistent environment that readers achieve reading fluency, which is

supported by one of the findings of this meta-analysis, i.e., repetitive practice produced larger effects than non-repetitive practice in reading rate, reading accuracy and comprehension. Extensive repetitive practice provides readers with successive exposure to print. Therefore, as letters, words, and texts or stories become increasingly familiar to readers through extensive and repeated practice, readers achieve automaticity in decoding or word recognition. Less attention is needed for these skills, freeing up more mental resources or attention for understanding the texts.

### *Implications for Future Research*

The summary of methodological and procedural information about the studies in the databases can serve as a useful resource for future research or evaluation studies. Cordray and Sonnefeld (1985) present a method for using meta-analytic information to plan an evaluation. For example, the average effect sizes for various intervention types might be used in a power analysis to determine the sample size sufficient for a proposed study.

As I pointed out at the beginning of this chapter, one of the major limitations of the current meta-analysis is the small number of quality studies included. It clearly shows that there is a need for more well-designed primary studies in reading fluency research. These studies need to include a control group so they can examine specific fluency building procedures contrasted with a control group and contrasted against each other. They also need to include reliable and valid outcome measures in fluency and comprehension. Only then is it possible that enough primary studies

examining various specific strategies will be available to allow meaningful comparisons made among various strategies and to allow practically significant conclusions drawn for teachers and practitioners. Further, the strategies or methods used should be characterized not by labels such as repeated reading, explicit teaching, direct instruction, etc., but by treatment descriptions that are explicit with regard to what kind of reading instruction it is, how much rereading there is, the nature and timing of the feedback, and the kind and level of materials. Changes that take place in student reading and knowledge should be documented during the intervention rather than just at the end. Only in this way would it be possible to clearly categorize fluency building strategies to make meaningful comparisons and arrive at practically significant conclusions.

There is also a need for longitudinal research that examines the impact of the various fluency building procedures on reading development of different types of readers at different points along the age continuum. Longitudinal studies of the impact of these procedures could clarify how long the intervention benefits can be maintained. It should be especially useful if these were examined with different learner types under various conditions in terms of type of passages, level of passage difficulty, feedback procedures, and outcome measure quality.

Related to the above point, there is also a need to for researchers to use more standard outcome measures of fluency and comprehension when remedial readers and students with disabilities are used as subjects. This way, the effect sizes will be meaningfully comparable with those obtained from studies involving normal readers.

APPENDIX A:  
Study Inclusion Criteria

## APPENDIX A STUDY INCLUSION CRITERIA

- A. There must be some intervention or treatment to increase reading fluency. Reading fluency includes reading rate, reading accuracy and reading comprehension (see definition below).
- B. The intervention must be conducted in the classrooms, at least in part, in public or private school settings, thus excluding intervention done by parents or at home. A school may be public or private. Preschools were not eligible. Interventions that included family components were eligible if that some portion of the program was implemented in a school, or through a school. After-school programs were eligible as long as they were implemented, at least in part, in a school setting, or implemented through a school.
- C. The study must assess intervention effects on children (defined below) for at least one outcome variable that represents oral reading fluency. Qualifying outcome variables specifically include those listed below:

Reading rate/speed: defined as words read per minute or number of seconds needed to read a passage of certain length. Reading rate could also be measured in reading age or grade equivalent.

Reading accuracy: defined as errors per minute or percent of errors, or reading age or grade equivalent.

Studies that measured oral reading fluency and reading comprehension were eligible, but those that measured only reading comprehension were not eligible.

Reading comprehension may include number of questions answered correctly, number of proposition recalled correctly, or performance on a standardized reading comprehension test.

- D. Quantitative data must be reported for at least one qualifying outcome variable measured on children. That is, the study must report original data, means, results of statistical tests, p values, ect. from which a treatment/control effect size can be calculated for oral reading fluency (and comprehension if comprehension is measured) related outcome variable.
- E. Subjects are elementary-aged children with or without special education needs. Elementary age children refer to those attending classes in grades K-6. If elementary age and older children were included in a study, the study was acceptable only if the results for the “elementary age” group were reported separately, or if the elementary age children constituted a majority of the subjects for whom results were reported, and it was report and it was plausible that the

oldest subjects in the sample were not older than 14. Subjects receiving special education services include those who are labeled as LD, ED/EBD, mild MR, or speech/language. Subjects without special education needs included students who had no reading problems and those were considered remedial readers. Subjects receiving remedial reading include those who are considered poor readers, below average readers, below grade readers, poorest readers, non-fluent readers, slow readers, etc.

- F. Research Design: Studies were acceptable only if they used experimental and quasi-experimental designs that compare subject groups receiving one or more identifiable treatments with one or more “control” conditions. Control conditions may be “no treatment,” “treatment as usual,” or any other similar condition set up as a contrast to the treatment condition that does not represent a concerted effort to produce change.

To be eligible as an experimental/quasi-experimental design, a study must meet at least one of the following criteria:

- a. Subjects were randomly assigned to treatment and control conditions or assigned by a procedure plausibly equivalent to randomization, e.g., arbitrary wait-list.
- b. Subjects in the treatment and control conditions were matched on pretest variables and/or other relevant personal and demographic characteristics.
- c. If subjects are not randomly assigned or matched, information about initial group differences on key variables must be presented in sufficient detail to permit coding of the degree of initial (pretreatment) equivalence. That is, the study must have both pretests and posttests on one or more outcome variables and/or a range of measures relevant to reading, such as IQ, reading achievement, etc which allows assessment of how similar the treatment and control groups are. Post-only non-equivalent comparisons (not randomized or matched) are not eligible.

Multiple group, multiple treatment studies that used experimental or quasi-experimental (i.e., randomized or matched) but did not include a control group as defined above are not eligible.

Pretest-posttest designs in which measures of one or more of the eligible outcome variables were taken before treatment and after treatment on the same group of subjects were not acceptable. These may be either of two types of designs:

- a. One group pretest-posttest studies: One subject group is pretested, receives treatment, and is posttested. Pre-post comparisons are available for at least one eligible outcome variable.
- b. Multiple group, multiple-treatment pre-post designs: More than one subject group is involved, each receiving a different treatment.

Post-only non-equivalent comparisons (not randomized or matched) are not eligible. Similarly, one-group or treatment-control designs that use the ABAB format (one type of single subject design) are not eligible.

G. The study is set in an English speaking country and is reported in English.

H. The date of publication is 1966 or later.



APPENDIX B:  
Study Level Coding Manual

APPENDIX B  
STUDY LEVEL CODING MANUAL

Bibliographic reference: Write a complete citation in (approx.) APA form.

1. Study ID Number. Assign a unique identification number to each study. If a report presents two independent studies, i.e., two independent outcome studies with different subjects, then add a decimal to the study ID number to distinguish each study within a report and code each independent study separately.
2. What type of publication is the report? If two separate reports are being used to code a single study, code the type of the more formally published report (i.e., book or journal article).

1 book	4 technical report
2 journal article or book chapter	5 conference paper
3 thesis or doctoral dissertation	6 other (specify)_____
3. What is the publication year (last two digits; 999 if unknown)? If two separate reports are being used to code a single study, code the publication year of the more formally published report.

**Sample Descriptors**

4. Mean age of sample. Specify the approximate or exact mean age at the beginning of intervention. Code the best information available; estimate mean age from grade levels if necessary. If mean age cannot be determined, enter "99.99."
5. Predominant race. Select the code that best describes the racial makeup of the sample.

1 greater than 60% White	5 mixed, none more than 60%
2 greater than 60% Black	6 mixed, cannot tell proportion
3 greater than 60% Hispanic	9 cannot tell
4 greater than 60% Other minority	
6. Predominant sex of sample. Select the code that best describes the proportion of males to the females in the sample.

1 less than 5% male	4 between 51% and 95% male
2 between 5% and 49% male	5 greater than 95% male
3 50% male	9 cannot tell
7. Select the code that best describes the predominate level of reading ability of subjects at onset of treatment.

- 0 subjects read at grade level
- 1 subjects read one year below grade level
- 2 subjects read two years below grade level
- 3 subjects read three years below grade level

8. Subject characteristic. Select the code that best describes the characteristic of the subjects.

- 1 students with disabilities (LD, ED/BD, mild MR, Speech/Language)
- 2 remedial readers (poor, poorest, below grade, below average, non-fluent)
- 3 normal readers (average, above average, or good readers)
- 9 cannot tell

9. Grade Level. Specify the grade levels of the subjects in the sample at onset of treatment.

- 1 first through third grade
- 2 fourth through sixth grade
- 9 cannot tell

#### Research Design Descriptors

10. Unit of assignment to conditions. Select the code that best describes the unit of assignment to treatment and control groups.

- 1 individual
- 2 pair
- 3 classroom/small group
- 9 cannot tell

11. Type of assignment to conditions. Select the code that best describes how subjects were assigned to treatment and control groups.

- 1 random, after matching, stratification, blocking, etc.
- 2 simple random (also includes systematic sampling)
- 3 nonrandom, post hoc matching
- 4 nonrandom, other
- 5 other (specify)
- 9 cannot tell

12. Overall confidence of judgment on how subjects were assigned.

- 1 very low (no basis)
- 2 low (guess)
- 3 moderate (weak inference)
- 4 high (strong inference)
- 5 very high (explicitly stated)

13. Was the equivalence of the groups tested at pretest?

- 1 yes
- 2 no

14. Pretest differences if tested. Note: an “important” difference means a difference on several variables or on a major variable, or large differences; major variables are those likely related to reading fluency and comprehension, e.g., IQ, sex, age, ethnicity, SES. Pretest differences on an outcome variable should be coded as “important.”

- 1 negligible differences, judged unimportant
- 2 some differences, judged of uncertain importance
- 3 some differences, judged important

15. Total sample size (start of study).

16. Control group sample size (start of study).

17. Treatment1 group sample size (start of study).

18. Treatment2 group sample size (start of study).

19. Treatment3 sample size (start of study).

20. Treatment4 sample size (start of study).

21. Treatment5 sample size (start of study)

22. Treatment6 sample size (start of study).

23. Treatment7 sample size (start of study).

### ***Nature of the Treatment Descriptors***

24. Type of treatment. Indicate the dominant treatment type of the program. Look for explicit mention of treatment technique in the report from description of treatment sessions. See program description for definitions of treatment. If the program uses more than one treatment type, indicate the one that appears most central to the treatment. If you cannot decide between two types, indicate both in the “other” category.

- 1 repeated reading
- 2 previewing
- 3 reading while listening
- 4 peer-tutoring
- 5 oral recitation lesson (ORL)
- 6 neurological impress method (NIM)
- 7 computer-assisted instruction
- 8 explicit teaching/effective teaching
- 9 language experience approach
- 10 decoding/word training
- 11 shared book experience

12 other (specify)

25. Recategorize the above treatment based on the following categories. Look for explicit mention of training content in the report from description of treatment sessions. See program description for specific content areas that students have been trained on. If you cannot decide between two types, indicate both in the "other" category.

1 repeated reading of sounds, words, and phrases.

2 repeated reading of stories/passages

3 repeated reading of sounds, words, phrases, and stories/passages

4 nonrepetitive reading of sounds, words, and phrases

5 non-repetitive reading of stories/passages

6 non-repetitive reading of sounds, words, phrases, and stories/passages.

7 other – specific classroom instructional methods

26. Format in which the treatment was conducted. Indicate how the intervention activities were carried out, e.g., adult-directed as a whole class or small group, tutoring, by an adult, peer, or cross-age, or subjects individually and independently engaged in interventions.

1 adult-directed, small group or whole class

2 tutoring, adults

3 tutoring, peers

4 tutoring, cross-age

5 individual independently

27. Treatment duration in weeks (missing=999). Approximate (or exact) duration of treatment in weeks from first treatment event to last treatment event, excluding follow-ups designated as such (divide # of days by 7 and round; multiple # of months by 4.3 and round). Estimate, if necessary.

28. Frequency of treatment. Indicate how often the treatment was conducted.

0 less than weekly

1 once a week

1 twice a week

2 three times a week

3 four times a week

5 daily

29. Length of each session, in minutes.

1 15 minutes

2 20 minutes

3 25 minutes

4 30 minutes

4 35 minutes

5 45 minutes

9 cannot tell

30. Were students trained to mastery/criterion in reading accuracy?

- 1 yes
- 2 no
- 9 cannot tell

31. If students were trained to mastery/criterion in accuracy, what criterion was used?

- 1 Error free/100% accuracy
- 2 95% accuracy
- 3 85% accuracy
- 9 cannot tell

32. Were students trained to mastery/criterion in reading rate?

- 1 yes
- 2 no
- 9 cannot tell

33. If students were trained to mastery/criterion in rate, what criterion was used? Indicate the number of words read per minute.

34. Was corrective feedback provided during intervention?

- 1 yes
- 2 no
- 9 cannot tell

35. Nature of control group.

- 1 receives nothing; no evidence of any treatment or attention; may still be in school.
- 2 treatment as usual, school setting; control receives the usual treatment in a school setting without the special enhancement that constitutes the treatment of interest; this refers to treatment occurring within a framework common to experimental and control groups with something added for the experimental group
- 3 attention placebo, e.g., control group receives discussion, attention, art, or deliberately diluted version of treatment.
- 4 alternative treatment; control is not really a control but another treatment (other than usual treatment) being compared with the focal treatment; only eligible if the alternative treatment is designed as a contrast and is not expected to work very well (straw man).
- 99 cannot tell

36. Overall confidence rating of judgment on the nature of the control group.

- 1 very low (no basis)
- 2 low (guess)
- 3 moderate (weak inference)

- 4 high (strong inference)
- 5 very high (explicitly stated)

37. Primary focus of treatment. Indicate whether the primary goal of the treatment was to increase reading fluency, or comprehension or both.

- 1 reading fluency only
- 2 comprehension only
- 3 fluency and comprehension
- 9 cannot tell

APPENDIX C:  
Effect Size Level Coding Manual



APPENDIX C  
EFFECT SIZE LEVEL CODING MANUAL

For each effect size, code all of the items below. Note that studies will have different numbers of effect sizes, and hence, different numbers of effect size level data coding forms.

1. Study ID number. Assign each study a unique identification number. If a report presents two independent studies, i.e., studies with different participants and generally a different design, then add a decimal to the study ID number to distinguish each study within a report.
2. Effect size number. Assign each effect size within a study a unique number. Number multiply effect sizes within a study sequentially, e.g., 1, 2, 3, 4, etc.

***Dependent Measure Descriptors***

3. Effect size type. Code an effect size as a pretest comparison if the measures being compared across groups were taken prior to the intervention, including reading rate, reading accuracy, and reading comprehension. Posttest effect sizes are the first reported comparison between the groups following the intervention for a construct, e.g. if reading rate at ten weeks following the intervention is the earliest post-treatment fluency measure, then it should be coded as a posttest comparison. Code any effect sizes for measures at future points in time as follow-up comparisons.
  - 1 pretest comparison
  - 2 posttest comparison
  - 3 pre-post gain score comparison
  - 4 follow-up comparison
4. Category of outcome construct.
  - 1 reading rate – speed/rate/number of words per minute
  - 2 reading accuracy - number of errors per minute, percent of errors
  - 3 reading comprehension

***Effect Size Data***

5. Type of data effect sizes based on
  - 1 means and standard deviations
  - 2 *t*-value or *F*-value
  - 3 chi-square (*df*=1)
  - 4 other \_\_\_\_\_
6. Page number where the data for this effect size were found.

## Calculating Effect Size

### ***When means and standard deviations are reported or can be estimated:***

7. Control group sample size (write in appropriate number).
8. Control group mean (write in the value for the mean, if available)
9. Control group standard deviation (write in the value for the sd, if available).

#### Treatment 1

10. Sample size (write in appropriate number).
11. Raw difference favors (i.e., shows more “success” for) which group?
  - 1 treatment group
  - 2 neither (exactly equal)
  - 3 control group
  - 4 treatment group, though statistically nonsignificant
  - 5 control group, though statistically nonsignificant
  - 9 cannot tell
12. Group mean (write in the value for the mean, if available).
  - a. Reading rate \_\_\_\_\_
  - b. Reading accuracy \_\_\_\_\_
  - c. Reading comprehension \_\_\_\_\_
13. Standard deviation (write in the value for the sd, if available).
  - a. Reading rate \_\_\_\_\_
  - b. Reading accuracy \_\_\_\_\_
  - c. Reading comprehension \_\_\_\_\_
14. Effect Size using the Excel Effect Size Determination Program or calculated by hand using appropriate procedures. Report to two decimals with an algebraic sign in font: plus if difference favors treatment; minus if difference favors control; +9.99 if NA.
  - a. ES for reading rate \_\_\_\_\_
  - b. ES for reading accuracy \_\_\_\_\_
  - c. ES for reading comprehension \_\_\_\_\_

#### Treatment 2

15. Sample size (write in appropriate number).

16. Raw difference favors (i.e., shows more “success” for) which group?
- 1 treatment group
  - 2 neither (exactly equal)
  - 3 control group
  - 4 treatment group, though statistically nonsignificant
  - 5 control group, though statistically nonsignificant
  - 9 cannot tell
17. Group mean (write in the value for the mean, if available).
- a. Reading rate \_\_\_\_\_
  - b. Reading accuracy \_\_\_\_\_
  - c. Reading comprehension \_\_\_\_\_
18. Standard deviation (write in the value for the sd, if available).
- a. Reading rate \_\_\_\_\_
  - b. Reading accuracy \_\_\_\_\_
  - c. Reading comprehension \_\_\_\_\_
19. Effect Size using the Excel Effect Size Determination Program or calculated by hand using appropriate procedures. Report to two decimals with an algebraic sign in font: plus if difference favors treatment; minus if difference favors control; +9.99 if NA.
- a. ES for reading rate \_\_\_\_\_
  - b. ES for reading accuracy \_\_\_\_\_
  - c. ES for reading comprehension \_\_\_\_\_

### Treatment 3

20. Sample size (write in appropriate number).
21. Raw difference favors (i.e., shows more “success” for) which group?
- 1 treatment group
  - 2 neither (exactly equal)
  - 3 control group
  - 4 treatment group, though statistically nonsignificant
  - 5 control group, though statistically nonsignificant
  - 9 cannot tell
22. Group mean (write in the value for the mean, if available).
- a. Reading rate \_\_\_\_\_
  - b. Reading accuracy \_\_\_\_\_
  - c. Reading comprehension \_\_\_\_\_

23. Standard deviation (write in the value for the sd, if available).

- a. Reading rate \_\_\_\_\_
- b. Reading accuracy \_\_\_\_\_
- c. Reading comprehension \_\_\_\_\_

24. Effect Size using the Excel Effect Size Determination Program or calculated by hand using appropriate procedures. Report to two decimals with an algebraic sign in font: plus if difference favors treatment; minus if difference favors control; +9.99 if NA.

- a. ES for reading rate \_\_\_\_\_
- b. ES for reading accuracy \_\_\_\_\_
- c. ES for reading comprehension \_\_\_\_\_

#### Treatment 4

25. Sample size (write in appropriate number).

26. Raw difference favors (i.e., shows more "success" for) which group?

- 1 treatment group
- 2 neither (exactly equal)
- 3 control group
- 4 treatment group, though statistically nonsignificant
- 5 control group, though statistically nonsignificant
- 9 cannot tell

27. Group mean (write in the value for the mean, if available).

- a. Reading rate \_\_\_\_\_
- b. Reading accuracy \_\_\_\_\_
- c. Reading comprehension \_\_\_\_\_

28. Standard deviation (write in the value for the sd, if available).

- a. Reading rate \_\_\_\_\_
- b. Reading accuracy \_\_\_\_\_
- c. Reading comprehension \_\_\_\_\_

29. Effect Size using the Excel Effect Size Determination Program or calculated by hand using appropriate procedures. Report to two decimals with an algebraic sign in font: plus if difference favors treatment; minus if difference favors control; +9.99 if NA.

- d. ES for reading rate \_\_\_\_\_
- e. ES for reading accuracy \_\_\_\_\_
- f. ES for reading comprehension \_\_\_\_\_

Treatment 5

30. Sample size (write in appropriate number).
31. Raw difference favors (i.e., shows more “success” for) which group?  
1 treatment group  
2 neither (exactly equal)  
3 control group  
4 treatment group, though statistically nonsignificant  
5 control group, though statistically nonsignificant  
9 cannot tell
32. Group mean (write in the value for the mean, if available).  
a. Reading rate \_\_\_\_\_  
b. Reading accuracy \_\_\_\_\_  
c. Reading comprehension \_\_\_\_\_
33. Standard deviation (write in the value for the sd, if available).  
a. Reading rate \_\_\_\_\_  
b. Reading accuracy \_\_\_\_\_  
c. Reading comprehension \_\_\_\_\_
34. Effect Size using the Excel Effect Size Determination Program or calculated by hand using appropriate procedures. Report to two decimals with an algebraic sign in font: plus if difference favors treatment; minus if difference favors control; +9.99 if NA.  
a. ES for reading rate \_\_\_\_\_  
b. ES for reading accuracy \_\_\_\_\_  
c. ES for reading comprehension \_\_\_\_\_

Treatment 6

35. Sample size (write in appropriate number).
36. Raw difference favors (i.e., shows more “success” for) which group?  
1 treatment group  
2 neither (exactly equal)  
3 control group  
4 treatment group, though statistically nonsignificant  
5 control group, though statistically nonsignificant  
9 cannot tell
37. Group mean (write in the value for the mean, if available).  
a. Reading rate \_\_\_\_\_

- b. Reading accuracy \_\_\_\_\_
  - c. Reading comprehension \_\_\_\_\_
38. Standard deviation (write in the value for the sd, if available).
- a. Reading rate \_\_\_\_\_
  - b. Reading accuracy \_\_\_\_\_
  - c. Reading comprehension \_\_\_\_\_
39. Effect Size using the Excel Effect Size Determination Program or calculated by hand using appropriate procedures. Report to two decimals with an algebraic sign in font: plus if difference favors treatment; minus if difference favors control; +9.99 if NA.
- a. ES for reading rate \_\_\_\_\_
  - b. ES for reading accuracy \_\_\_\_\_
  - c. ES for reading comprehension \_\_\_\_\_

Treatment 7

40. Sample size (write in appropriate number).
41. Raw difference favors (i.e., shows more “success” for) which group?
- 1 treatment group
  - 2 neither (exactly equal)
  - 3 control group
  - 4 treatment group, though statistically nonsignificant
  - 5 control group, though statistically nonsignificant
  - 9 cannot tell
42. Group mean (write in the value for the mean, if available).
- a. Reading rate \_\_\_\_\_
  - b. Reading accuracy \_\_\_\_\_
  - c. Reading comprehension \_\_\_\_\_
43. Standard deviation (write in the value for the sd, if available).
- a. Reading rate \_\_\_\_\_
  - b. Reading accuracy \_\_\_\_\_
  - c. Reading comprehension \_\_\_\_\_
44. Effect Size using the Excel Effect Size Determination Program or calculated by hand using appropriate procedures. Report to two decimals with an algebraic sign in font: plus if difference favors treatment; minus if difference favors control; +9.99 if NA.
- a. ES for reading rate \_\_\_\_\_
  - b. ES for reading accuracy \_\_\_\_\_
  - c. ES for reading comprehension \_\_\_\_\_

***When significant test information is reported:***

45. *t*-value (write in the value, if available).
46. *F*-value (write in the value, if available).
47. Confidence rating in effect size computation.
  - 1 highly estimated (have *N* and crude *p* value only, such as  $p < .10$ , and must reconstruct via rough *t*-test equivalence)
  - 2 moderate estimation (have complex but relatively complete statistics, such as multifactor ANOVA, as basis for estimation)
  - 3 some estimation (have unconventional statistics and must convert to equivalent *t*-values or have conventional statistics but incomplete, such as exact *p* level)
  - 4 slight estimation (must use significance testing statistics rather than descriptive statistics, but have complete statistics of conventional sort)
  - 5 no estimation (have descriptive data such as means, standard deviations, frequencies, proportions, etc. and can calculate effect size directly)

APPENDIX D:  
Study Level Coding Form



## APPENDIX D STUDY LEVEL CODING FORM

### ***Bibliographic Reference:***

- \_\_\_\_\_ 1. Study ID Number [STUDYID]
- \_\_\_\_\_ 2. Type of publication[PUBTYPE]
- |                                    |                          |
|------------------------------------|--------------------------|
| 1. Book                            | 4. Technical report      |
| 2. Journal article or book chapter | 5. Conference paper      |
| 3. Thesis or doctoral dissertation | 6. Other (specify):_____ |
- \_\_\_\_\_ 3. Publication year (last two digits; 99 if unknown) [PUBYEAR]

### ***Sample Descriptors***

- \_\_\_\_\_ 4. Mean age [MEANAGE]
- \_\_\_\_\_ 5. Predominant race [RACE]
- |                       |                                     |
|-----------------------|-------------------------------------|
| 1 >60% White          | 5 mixed, none more than 60%         |
| 2 >60% Black          | 6 mixed, cannot estimate proportion |
| 3 >60% Hispanic       | 9 cannot tell                       |
| 4 >60% other minority |                                     |
- \_\_\_\_\_ 6. Predominant sex [SEX]
- |               |                |
|---------------|----------------|
| 1 <5% male    | 4 51%-95% male |
| 2 5%-49% male | 5 >95% male    |
| 3 50% male    | 9 cannot tell  |
- \_\_\_\_\_ 7. Reading level (RD\_LEVEL)
- |   |               |
|---|---------------|
| 0 subjects read at grade level                |               |
| 1 subjects read one year below grade level    | 9 cannot tell |
| 2 subjects read two years below grade level   |               |
| 3 subjects read three years below grade level |               |
- \_\_\_\_\_ 8. Subject characteristics.
- |  |  |
|--|--|
| 1 students with disabilities (LD, ED/BD, mild MR, Speech/Language)                 |  |
| 2 remedial readers (poorest, poor readers, below grade, below average, non-fluent) |  |
| 3 normal readers (average, above average, or good readers)                         |  |
| 9 cannot tell.   |  |
- \_\_\_\_\_ 9. Grade level [GRADE]
- |                              |  |
|------------------------------|--|
| 1 first through third grade  |  |
| 2 fourth through sixth grade |  |
| 9 cannot tell                |  |

### ***Research Design Descriptors***



- 2 twice a week
- 3 three times a week
- 4 four times a week
- 5 daily

\_\_\_\_\_ 21. Length of each training session, in minutes. [LENGTH]

- 1 10-20 minuts
- 2 21-30 minutes
- 3 31-40 minutes
- 4 41-50 minutes
- 5 more than 50 minutes
- 9 cannot tell

22. How were training materials determined?

- 1 based on students interest and relevance to their background
- 2 based on students reading level/readability
- 3 based on researcher decision/interest
- 4 based on availability
- 9 cannot tell

***Nature of Treatment Descriptors***

Treatment 1

\_\_\_\_\_ 23. Sample size (start of study). [ORIGN\_TX1]

\_\_\_\_\_ 24. Treatment type [TX1\_TYPE]

- 1 repeated reading
- 2 previewing
- 3 read while listening
- 4 oral recitation lesson
- 5 neurological impress method
- 6 computer-assisted instruction
- 7 language experience approach
- 8 decoding/word traing
- 9 shared book experience
- 10 other \_\_\_\_\_

\_\_\_\_\_ 25. Recategorize the above treatments based on the following categories. [TX1\_CATEG]

- 1 repeated reading of sounds, words, and phrases.
- 2 repeated reading of stories/passages.
- 3 repeated reading of sounds, words, phrases, and stories/passages.
- 4 non-repetitive reading of sounds, words, and phrases.
- 5 non-repetitive reading of stories/passages.
- 6 non-repetitive reading of sounds, words, phrases, and stories/passages.
- 7 other - specific classroom instructional methods

\_\_\_\_\_ 26. Format in which treatment was conducted. [TX1\_FMAT]

- 1 adult-directed, small group or whole class.
- 2 tutoring, adults.
- 3 tutoring, peers.
- 4 tutoring, cross-age
- 5 subject independently completed treatment (with supervision of adults)

- \_\_\_\_\_ 27. Were students trained to mastery/criterion in reading accuracy?  
 1 yes  
 2 no  
 9 cannot tell
- \_\_\_\_\_ 28. If students were trained to mastery/criterion in accuracy, what criterion was used?  
 1 error free/100% accuracy  
 2 95% accurate or better but less than 100% accurate  
 3 90% accuracy or better but less than 95% accurate  
 4 85% accurate or better but less than 90%  
 9 cannot tell
- \_\_\_\_\_ 29. Were students trained to mastery/criterion in reading rate?  
 1 yes  
 2 no  
 9 cannot tell
- \_\_\_\_\_ 30. If students were trained to mastery/criterion in reading rate, what criterion was used?
- \_\_\_\_\_ 31. Was corrective feedback provided during intervention?  
 1 yes  
 2 no  
 9 cannot tell
- \_\_\_\_\_ 32. Treatment focus. [TX\_FOCUS]  
 1 reading fluency only  
 2 comprehension only  
 3 fluency and comprehension  
 9 cannot tell
- \_\_\_\_\_ 33. What comprehension activities were involved in the intervention?  
 1 Oral discussion of stories  
 2 Retelling of stories  
 3 Asking and answering questions about stories (who, where, what, how)  
 4 writing about stories (comparisons, questions, sequences, summaries, etc.)  
 9 cannot tell

Treatment 2

- \_\_\_\_\_ 34. Sample size (start of study). [ORIGN\_TX1]
- \_\_\_\_\_ 35. Treatment type [TX1\_TYPE]  
 1 repeated reading  
 2 previewing  
 3 read while listening  
 4 oral recitation lesson  
 5 neurological impress method  
 6 computer-assisted instruction  
 7 language experience approach  
 8 decoding/word traing  
 9 shared book experience  
 10 other\_\_\_\_\_
- \_\_\_\_\_ 36. Recategorize the above treatments based on the following categories. [TX1\_CATEG]

- 1 repeated reading of sounds, words, and phrases.
- 2 repeated reading of stories/passages.
- 3 repeated reading of sounds, words, phrases, and stories/passages.
- 4 non-repetitive reading of sounds, words, and phrases.
- 5 non-repetitive reading of stories/passages.
- 6 non-repetitive reading of sounds, words, phrases, and stories/passages.
- 7 other - specific classroom instructional methods

\_\_\_\_\_ 37. Format in which treatment was conducted. [TX1\_FMAT]

- 1 adult-directed, small group or whole class.
- 2 tutoring, adults.
- 3 tutoring, peers.
- 4 tutoring, cross-age
- 5 subject independently completed treatment (with supervision of adults)

\_\_\_\_\_ 38. Were students trained to mastery/criterion in reading accuracy?

- 1 yes
- 2 no
- 9 cannot tell

\_\_\_\_\_ 39. If students were trained to mastery/criterion in accuracy, what criterion was used?

- 1 error free/100% accuracy
- 2 95% accurate or better but less than 100% accurate
- 3 90% accuracy or better but less than 95% accurate
- 4 85% accurate or better but less than 90%
- 9 cannot tell

\_\_\_\_\_ 40. Were students trained to mastery/criterion in reading rate?

- 1 yes
- 2 no
- 9 cannot tell

\_\_\_\_\_ 41. If students were trained to mastery/criterion in reading rate, what criterion was used?

\_\_\_\_\_ 42. Was corrective feedback provided during intervention?

- 1 yes
- 2 no
- 9 cannot tell

\_\_\_\_\_ 43. Treatment focus. [TX\_FOCUS]

- 1 reading fluency only
- 2 comprehension only
- 3 fluency and comprehension
- 9 cannot tell

\_\_\_\_\_ 44. What comprehension activities were involved in the intervention?

- 1 Oral discussion of stories
- 2 Retelling of stories
- 3 Asking and answering questions about stories (who, where, what, how)
- 4 writing about stories (comparisons, questions, sequences, summaries, etc.)
- 9 cannot tell

Treatment 3

- \_\_\_\_\_ 45. Sample size (start of study). [ORIGN\_TX1]
- \_\_\_\_\_ 46. Treatment type [TX1\_TYPE]
- 1 repeated reading
  - 2 previewing
  - 3 read while listening
  - 4 oral recitation lesson
  - 5 neurological impress method
  - 6 computer-assisted instruction
  - 7 language experience approach
  - 8 decoding/word traing
  - 9 shared book experience
  - 10 other\_\_\_\_\_
- \_\_\_\_\_ 47. Recategorize the above treatments based on the following categories. [TX1\_CATEG]
- 1 repeated reading of sounds, words, and phrases.
  - 2 repeated reading of stories/passages.
  - 3 repeated reading of sounds, words, phrases, and stories/passages.
  - 4 non-repetitive reading of sounds, words, and phrases.
  - 5 non-repetitive reading of stories/passages.
  - 6 non-repetitive reading of sounds, words, phrases, and stories/passages.
  - 7 other - specific classroom instructional methods
- \_\_\_\_\_ 48. Format in which treatment was conducted. [TX1\_FMAT]
- 1 adult-directed, small group or whole class.
  - 2 tutoring, adults.
  - 3 tutoring, peers.
  - 4 tutoring, cross-age
  - 5 subject independently completed treatment (with supervision of adults)
- \_\_\_\_\_ 49. Were students trained to mastery/criterion in reading accuracy?
- 1 yes
  - 2 no
  - 9 cannot tell
- \_\_\_\_\_ 50. If students were trained to mastery/criterion in accuracy, what criterion was used?
- 1 error free/100% accuracy
  - 2 95% accurate or better but less than 100% accurate
  - 3 90% accuracy or better but less than 95% accurate
  - 4 85% accurate or better but less than 90%
  - 9 cannot tell
- \_\_\_\_\_ 51. Were students trained to mastery/criterion in reading rate?
- 1 yes
  - 2 no
  - 9 cannot tell
- \_\_\_\_\_ 52. If students were trained to mastery/criterion in reading rate, what criterion was used?
- \_\_\_\_\_ 53. Was corrective feedback provided during intervention?
- 1 yes
  - 2 no
  - 9 cannot tell

- \_\_\_\_\_ 54. Treatment focus. [TX\_FOCUS]  
 1 reading fluency only  
 2 comprehension only  
 3 fluency and comprehension  
 9 cannot tell
- \_\_\_\_\_ 55. What comprehension activities were involved in the intervention?  
 1 Oral discussion of stories  
 2 Retelling of stories  
 3 Asking and answering questions about stories (who, where, what, how)  
 4 writing about stories (comparisons, questions, sequences, summaries, etc.)  
 9 cannot tell

Treatment 4

- \_\_\_\_\_ 56. Sample size (start of study). [ORIGN\_TX1]
- \_\_\_\_\_ 57. Treatment type [TX1\_TYPE]  
 1 repeated reading  
 2 previewing  
 3 read while listening  
 4 oral recitation lesson  
 5 neurological impress method  
 6 computer-assisted instruction  
 7 language experience approach  
 8 decoding/word traing  
 9 shared book experience  
 10 other \_\_\_\_\_
- \_\_\_\_\_ 58. Recategorize the above treatments based on the following categories. [TX1\_CATEG]  
 1 repeated reading of sounds, words, and phrases.  
 2 repeated reading of stories/passages.  
 3 repeated reading of sounds, words, phrases, and stories/passages.  
 4 non-repetitive reading of sounds, words, and phrases.  
 5 non-repetitive reading of stories/passages.  
 6 non-repetitive reading of sounds, words, phrases, and stories/passages.  
 7 other - specific classroom instructional methods
- \_\_\_\_\_ 59. Format in which treatment was conducted. [TX1\_FMAT]  
 1 adult-directed, small group or whole class.  
 2 tutoring, adults.  
 3 tutoring, peers.  
 4 tutoring, cross-age  
 5 subject independently completed treatment (with supervision of adults)
- \_\_\_\_\_ 60. Were students trained to mastery/criterion in reading accuracy?  
 1 yes  
 2 no  
 9 cannot tell
- \_\_\_\_\_ 61. If students were trained to mastery/criterion in accuracy, what criterion was used?  
 1 error free/100% accuracy  
 2 95% accurate or better but less than 100% accurate  
 3 90% accuracy or better but less than 95% accurate

- 4 85% accurate or better but less than 90%
- 9 cannot tell

\_\_\_\_\_ 62. Were students trained to mastery/criterion in reading rate?

- 1 yes
- 2 no
- 9 cannot tell

\_\_\_\_\_ 63. If students were trained to mastery/criterion in reading rate, what criterion was used?

\_\_\_\_\_ 64. Was corrective feedback provided during intervention?

- 1 yes
- 2 no
- 9 cannot tell

\_\_\_\_\_ 65. Treatment focus. [TX\_FOCUS]

- 1 reading fluency only
- 2 comprehension only
- 3 fluency and comprehension
- 9 cannot tell

\_\_\_\_\_ 66. What comprehension activities were involved in the intervention?

- 1 Oral discussion of stories
- 2 Retelling of stories
- 3 Asking and answering questions about stories (who, where, what, how)
- 4 writing about stories (comparisons, questions, sequences, summaries, etc.)
- 9 cannot tell

#### Treatment 5

\_\_\_\_\_ 67. Sample size (start of study). [ORIGN\_TX1]

\_\_\_\_\_ 68. Treatment type [TX1\_TYPE]

- 1 repeated reading
- 2 previewing
- 3 read while listening
- 4 oral recitation lesson
- 5 neurological impress method
- 6 computer-assisted instruction
- 7 language experience approach
- 8 decoding/word traing
- 9 shared book experience
- 10 other \_\_\_\_\_

\_\_\_\_\_ 69. Recategorize the above treatments based on the following categories. [TX1\_CATEG]

- 1 repeated reading of sounds, words, and phrases.
- 2 repeated reading of stories/passages.
- 3 repeated reading of sounds, words, phrases, and stories/passages.
- 4 non-repetitive reading of sounds, words, and phrases.
- 5 non-repetitive reading of stories/passages.
- 6 non-repetitive reading of sounds, words, phrases, and stories/passages.
- 7 other - specific classroom instructional methods

\_\_\_\_\_ 70. Format in which treatment was conducted. [TX1\_FMAT]



- 1 adult-directed, small group or whole class.
- 2 tutoring, adults.
- 3 tutoring, peers.
- 4 tutoring, cross-age
- 5 subject independently completed treatment (with supervision of adults)

\_\_\_\_\_ 71. Were students trained to mastery/criterion in reading accuracy?  
 1 yes  
 2 no  
 9 cannot tell

\_\_\_\_\_ 72. If students were trained to mastery/criterion in accuracy, what criterion was used?  
 1 error free/100% accuracy  
 2 95% accurate or better but less than 100% accurate  
 3 90% accuracy or better but less than 95% accurate  
 4 85% accurate or better but less than 90%  
 9 cannot tell

\_\_\_\_\_ 73. Were students trained to mastery/criterion in reading rate?  
 1 yes  
 2 no  
 9 cannot tell

\_\_\_\_\_ 74. If students were trained to mastery/criterion in reading rate, what criterion was used?

\_\_\_\_\_ 75. Was corrective feedback provided during intervention?  
 1 yes  
 2 no  
 9 cannot tell

\_\_\_\_\_ 76. Treatment focus. [TX\_FOCUS]  
 1 reading fluency only  
 2 comprehension only  
 3 fluency and comprehension  
 9 cannot tell

\_\_\_\_\_ 77. What comprehension activities were involved in the intervention?  
 1 Oral discussion of stories  
 2 Retelling of stories  
 3 Asking and answering questions about stories (who, where, what, how)  
 4 writing about stories (comparisons, questions, sequences, summaries, etc.)  
 9 cannot tell

Treatment 6

\_\_\_\_\_ 78. Sample size (start of study). [ORIGN\_TX1]

\_\_\_\_\_ 79. Treatment type [TX1\_TYPE]  
 1 repeated reading  
 2 previewing  
 3 read while listening  
 4 oral recitation lesson  
 5 neurological impress method  
 6 computer-assisted instruction

- 7 language experience approach
- 8 decoding/word training
- 9 shared book experience
- 10 other \_\_\_\_\_

- \_\_\_\_\_ 80. Recategorize the above treatments based on the following categories. [TX1\_CATEG]
- 1 repeated reading of sounds, words, and phrases.
  - 2 repeated reading of stories/passages.
  - 3 repeated reading of sounds, words, phrases, and stories/passages.
  - 4 non-repetitive reading of sounds, words, and phrases.
  - 5 non-repetitive reading of stories/passages.
  - 6 non-repetitive reading of sounds, words, phrases, and stories/passages.
  - 7 other - specific classroom instructional methods
- \_\_\_\_\_ 81. Format in which treatment was conducted. [TX1\_FMAT]
- 1 adult-directed, small group or whole class.
  - 2 tutoring, adults.
  - 3 tutoring, peers.
  - 4 tutoring, cross-age
  - 5 subject independently completed treatment (with supervision of adults)
- \_\_\_\_\_ 82. Were students trained to mastery/criterion in reading accuracy?
- 1 yes
  - 2 no
  - 9 cannot tell
- \_\_\_\_\_ 83. If students were trained to mastery/criterion in accuracy, what criterion was used?
- 1 error free/100% accuracy
  - 2 95% accurate or better but less than 100% accurate
  - 3 90% accuracy or better but less than 95% accurate
  - 4 85% accurate or better but less than 90%
  - 9 cannot tell
- \_\_\_\_\_ 84. Were students trained to mastery/criterion in reading rate?
- 1 yes
  - 2 no
  - 9 cannot tell
- \_\_\_\_\_ 85. If students were trained to mastery/criterion in reading rate, what criterion was used?
- \_\_\_\_\_ 86. Was corrective feedback provided during intervention?
- 1 yes
  - 2 no
  - 9 cannot tell
- \_\_\_\_\_ 87. Treatment focus. [TX\_FOCUS]
- 1 reading fluency only
  - 2 comprehension only
  - 3 fluency and comprehension
  - 9 cannot tell
- \_\_\_\_\_ 88. What comprehension activities were involved in the intervention?
- 1 Oral discussion of stories
  - 2 Retelling of stories

- 3 Asking and answering questions about stories (who, where, what, how)
- 4 writing about stories (comparisons, questions, sequences, summaries, etc.)
- 9 cannot tell

Treatment 7

- \_\_\_\_\_ 89. Sample size (start of study). [ORIGN\_TX1]
  
- \_\_\_\_\_ 90. Treatment type [TX1\_TYPE]
  - 1 repeated reading
  - 2 previewing
  - 3 read while listening
  - 4 oral recitation lesson
  - 5 neurological impress method
  - 6 computer-assisted instruction
  - 7 language experience approach
  - 8 decoding/word traing
  - 9 shared book experience
  - 10 other\_\_\_\_\_
  
- \_\_\_\_\_ 91. Recategorize the above treatments based on the following categories. [TX1\_CATEG]
  - 1 repeated reading of sounds, words, and phrases.
  - 2 repeated reading of stories/passages.
  - 3 repeated reading of sounds, words, phrases, and stories/passages.
  - 4 non-repetitive reading of sounds, words, and phrases.
  - 5 non-repetitive reading of stories/passages.
  - 6 non-repetitive reading of sounds, words, phrases, and stories/passages.
  - 7 other - specific classroom instructional methods
  
- \_\_\_\_\_ 92. Format in which treatment was conducted. [TX1\_FMAT]
  - 1 adult-directed, small group or whole class.
  - 2 tutoring, adults.
  - 3 tutoring, peers.
  - 4 tutoring, cross-age
  - 5 subject independently completed treatment (with supervision of adults)
  
- \_\_\_\_\_ 93. Were students trained to mastery/criterion in reading accuracy?
  - 1 yes
  - 2 no
  - 9 cannot tell
  
- \_\_\_\_\_ 94. If students were trained to mastery/criterion in accuracy, what criterion was used?
  - 1 error free/100% accuracy
  - 2 95% accurate or better but less than 100% accurate
  - 3 90% accuracy or better but less than 95% accurate
  - 4 85% accurate or better but less than 90%
  - 9 cannot tell
  
- \_\_\_\_\_ 95. Were students trained to mastery/criterion in reading rate?
  - 1 yes
  - 2 no
  - 9 cannot tell
  
- \_\_\_\_\_ 96. If students were trained to mastery/criterion in reading rate, what criterion was used?

- \_\_\_\_\_ 97. Was corrective feedback provided during intervention?  
1 yes  
2 no  
9 cannot tell
- \_\_\_\_\_ 98. Treatment focus. [TX\_FOCUS]  
1 reading fluency only  
2 comprehension only  
3 fluency and comprehension  
9 cannot tell
- \_\_\_\_\_ 99. What comprehension activities were involved in the intervention?  
1 Oral discussion of stories  
2 Retelling of stories  
3 Asking and answering questions about stories (who, where, what, how)  
4 writing about stories (comparisons, questions, sequences, summaries, etc.)  
9 cannot tell

APPENDIX E:  
Effect Size Level Coding Form

APPENDIX E  
EFFECT SIZE LEVEL CODING FORM

\_\_\_\_\_ 1. Study ID number [STUDYID]

\_\_\_\_\_ 2. Effect size sequence number [ESNUM]

**Dependent Measure Descriptors**

\_\_\_\_\_ 3. Effect size type [ESTYPE]

- 1 pre-posttest comparison
- 2 post comparison
- 3 follow up comparison

\_\_\_\_\_ 4. Category of outcome construct [OUTCOME]

- 1 oral reading fluency-speed/rate/number of words per minute
- 2 oral reading fluency--accuracy or number of errors per minute
- 3 reading comprehension

**Calculating Effect Size**

\_\_\_\_\_ 5. Type of data effect size based on [DAT\_TYPE]

- 1 means and standard deviations
- 2 t-value or F-value
- 3 chi-square (df=1)
- 4 other \_\_\_\_\_

\_\_\_\_\_ 6. Page number where effect size data found [PAGENUM]

**When means and standard deviations are reported or can be estimated:**

\_\_\_\_\_ 7. Control group sample size [CN\_N]

\_\_\_\_\_ 8. Control group pretest mean

- \_\_\_\_\_ a. reading rate group mean [CN\_PRERATE]
- \_\_\_\_\_ b. reading accuracy group mean [CN\_PREACC]
- \_\_\_\_\_ c. reading comprehension group mean [CN\_PRECOMP]

\_\_\_\_\_ 9. Control group pretest standard deviation

- \_\_\_\_\_ a. reading rate standard deviation [CNS\_PRERATE]
- \_\_\_\_\_ b. reading accuracy standard deviation [CNS\_PREACC]
- \_\_\_\_\_ c. reading comprehension standard deviation [CNS\_PRECOMP]

\_\_\_\_\_ 10. Control group posttest mean

- \_\_\_\_\_ a. reading rate group mean [CN\_PSTRATE]
- \_\_\_\_\_ b. reading accuracy group mean [CN\_PSTACC]
- \_\_\_\_\_ c. reading comprehension group mean [CN\_PSTCOMP]

\_\_\_\_\_ 11. Control group pretest standard deviation

- \_\_\_\_\_ a. reading rate standard deviation [CNS\_PSTRATE]
- \_\_\_\_\_ b. reading accuracy standard deviation [CNS\_PSTACC]

\_\_\_\_\_

c. reading comprehension standard deviation [CNS\_PSTCOMP]

Treatment 1

\_\_\_\_\_

12. Sample size [TX1\_N]

\_\_\_\_\_

13. Group pretest mean

\_\_\_\_\_

a. reading rate [PRERATE1]

\_\_\_\_\_

b. accuracy [PREACC1]

\_\_\_\_\_

c. comprehension [PRECOMP1]

\_\_\_\_\_

14. Treatment group posttest standard deviation

\_\_\_\_\_

a. reading rate [ST1\_PSTRATE]

\_\_\_\_\_

b. reading accuracy [ST1\_PSTACC]

\_\_\_\_\_

c. comprehension [ST1\_PSTCOMP]

\_\_\_\_\_

15. Group posttest mean

\_\_\_\_\_

a. reading rate [PSTRATE1]

\_\_\_\_\_

b. accuracy [PSTACC1]

\_\_\_\_\_

c. comprehension [PSTCOMP1]

\_\_\_\_\_

16. Group posttest standard deviation

\_\_\_\_\_

a. reading rate [ST1\_PSTRATE]

\_\_\_\_\_

b. reading accuracy [ST1\_PSTACC]

\_\_\_\_\_

c. comprehension [ST1\_PSTCOMP]

\_\_\_\_\_

17. Raw difference favors (i.e., shows more success for) [TX1\_SUCS]

1 treatment group

2 neither (exactly equal)

3 control group

4 treatment group, though statistically nonsignificant

9 cannot tell

**When significant test information is reported:**

\_\_\_\_\_

18. t-value (write in the value if available)

\_\_\_\_\_

a. t-value for reading rate [T\_RATE1]

\_\_\_\_\_

b. t-value for accuracy [T\_ACC1]

\_\_\_\_\_

c. t-value for comprehension [T\_COMP1]

\_\_\_\_\_

19. F-value (write in the value if available)

\_\_\_\_\_

a. F-value for reading rate [F\_RATE1]

\_\_\_\_\_

b. F-value for accuracy [F\_ACC1]

\_\_\_\_\_

c. t-value for comprehension [F\_COMP1]

\_\_\_\_\_

20. Pretest Effect size for:

\_\_\_\_\_

a. reading rate [ES1\_PRERATE]

\_\_\_\_\_

b. reading accuracy [ES1\_PREACC]

\_\_\_\_\_

c. reading comprehension [ES1\_PRECOMP]

\_\_\_\_\_

21. Posttest Effect size for:

a. reading rate [ES1\_RATE]

b. reading accuracy [ES1\_ACC]

c. reading comprehension [ES1\_COMP]

Treatment 2

- \_\_\_\_\_ 22. Sample size [TX2\_N]
- \_\_\_\_\_ 23. Group pretest mean  
\_\_\_\_\_ a. reading rate [PRERATE2]  
\_\_\_\_\_ b. accuracy [PREACC2]  
\_\_\_\_\_ c. comprehension [PRECOMP2]
- \_\_\_\_\_ 24. Treatment group posttest standard deviation  
\_\_\_\_\_ a. reading rate [ST2\_PSTRATE]  
\_\_\_\_\_ b. reading accuracy [ST2\_PSTACC]  
\_\_\_\_\_ c. comprehension [ST2\_PSTCOMP]
- \_\_\_\_\_ 25. Group posttest mean  
\_\_\_\_\_ a. reading rate [PSTRATE2]  
\_\_\_\_\_ b. accuracy [PSTACC2]  
\_\_\_\_\_ c. comprehension [PSTCOMP2]
- \_\_\_\_\_ 26. Group posttest standard deviation  
\_\_\_\_\_ a. reading rate [ST2\_PSTRATE]  
\_\_\_\_\_ b. reading accuracy [ST2\_PSTACC]  
\_\_\_\_\_ c. comprehension [ST2\_PSTCOMP]
- \_\_\_\_\_ 27. Raw difference favors (i.e., shows more success for) [TX2\_SUCS]  
1 treatment group  
2 neither (exactly equal)  
3 control group  
4 treatment group, though statistically nonsignificant  
9 cannot tell

**When significant test information is reported:**

- \_\_\_\_\_ 28. t-value (write in the value if available)  
\_\_\_\_\_ a. t-value for reading rate [T\_RATE2]  
\_\_\_\_\_ b. t-value for accuracy [T\_ACC2]  
\_\_\_\_\_ c. t-value for comprehension [T\_COMP2]
- \_\_\_\_\_ 29. F-value (write in the value if available)  
\_\_\_\_\_ a. F-value for reading rate [F\_RATE2]  
\_\_\_\_\_ b. F-value for accuracy [F\_ACC2]  
\_\_\_\_\_ c. t-value for comprehension [F\_COMP2]
- \_\_\_\_\_ 30. Pretest Effect size for:  
\_\_\_\_\_ a. reading rate [ES2\_PRERATE]  
\_\_\_\_\_ b. reading accuracy [ES2\_PREACC]  
\_\_\_\_\_ c. reading comprehension [ES2\_PRECOMP]
- \_\_\_\_\_ 31. Posttest Effect size for:  
\_\_\_\_\_ a. reading rate [ES2\_RATE]  
\_\_\_\_\_ b. reading accuracy [ES2\_ACC]  
\_\_\_\_\_ c. reading comprehension [ES2\_COMP]



Treatment 3

- \_\_\_\_\_ 32. Sample size [TX3\_N]
- \_\_\_\_\_ 33. Group pretest mean  
\_\_\_\_\_ a. reading rate [PRERATE3]  
\_\_\_\_\_ b. accuracy [PREACC3]  
\_\_\_\_\_ c. comprehension [PRECOMP3]
- \_\_\_\_\_ 34. Treatment group posttest standard deviation  
\_\_\_\_\_ a. reading rate [ST3\_PSTRATE]  
\_\_\_\_\_ b. reading accuracy [ST3\_PSTACC]  
\_\_\_\_\_ c. comprehension [ST3\_PSTCOMP]
- \_\_\_\_\_ 35. Group posttest mean  
\_\_\_\_\_ a. reading rate [PSTRATE3]  
\_\_\_\_\_ b. accuracy [PSTACC3]  
\_\_\_\_\_ c. comprehension [PSTCOMP3]
- \_\_\_\_\_ 36. Group posttest standard deviation  
\_\_\_\_\_ a. reading rate [ST3\_PSTRATE]  
\_\_\_\_\_ b. reading accuracy [ST3\_PSTACC]  
\_\_\_\_\_ c. comprehension [ST3\_PSTCOMP]
- \_\_\_\_\_ 37. Raw difference favors (i.e., shows more success for) [TX3\_SUCS]  
1 treatment group  
2 neither (exactly equal)  
3 control group  
4 treatment group, though statistically nonsignificant  
9 cannot tell

**When significant test information is reported:**

- \_\_\_\_\_ 38. t-value (write in the value if available)  
\_\_\_\_\_ a. t-value for reading rate [T\_RATE3]  
\_\_\_\_\_ b. t-value for accuracy [T\_ACC3]  
\_\_\_\_\_ c. t-value for comprehension [T\_COMP3]
- \_\_\_\_\_ 39. F-value (write in the value if available)  
\_\_\_\_\_ a. F-value for reading rate [F\_RATE3]  
\_\_\_\_\_ b. F-value for accuracy [F\_ACC3]  
\_\_\_\_\_ c. t-value for comprehension [F\_COMP3]
- \_\_\_\_\_ 40. Pretest Effect size for:  
\_\_\_\_\_ a. reading rate [ES3\_PRERATE]  
\_\_\_\_\_ b. reading accuracy [ES3\_PREACC]  
\_\_\_\_\_ c. reading comprehension [ES3\_PRECOMP]
- \_\_\_\_\_ 41. Posttest Effect size for:  
\_\_\_\_\_ a. reading rate [ES3\_RATE]  
\_\_\_\_\_ b. reading accuracy [ES3\_ACC]  
\_\_\_\_\_ c. reading comprehension [ES3\_COMP]

Treatment 4

- \_\_\_\_\_
42. Sample size [TX4\_N]
- \_\_\_\_\_
43. Group pretest mean
- \_\_\_\_\_ a. reading rate [PRERATE4]
- \_\_\_\_\_ b. accuracy [PREACC4]
- \_\_\_\_\_ c. comprehension [PRECOMP4]
- \_\_\_\_\_
44. Treatment group posttest standard deviation
- \_\_\_\_\_ a. reading rate [ST4\_PSTRATE]
- \_\_\_\_\_ b. reading accuracy [ST4\_PSTACC]
- \_\_\_\_\_ c. comprehension [ST4\_PSTCOMP]
- \_\_\_\_\_
45. Group posttest mean
- \_\_\_\_\_ a. reading rate [PSTRATE4]
- \_\_\_\_\_ b. accuracy [PSTACC4]
- \_\_\_\_\_ c. comprehension [PSTCOMP4]
- \_\_\_\_\_
46. Group posttest standard deviation
- \_\_\_\_\_ a. reading rate [ST4\_PSTRATE]
- \_\_\_\_\_ b. reading accuracy [ST4\_PSTACC]
- \_\_\_\_\_ c. comprehension [ST4\_PSTCOMP]
- \_\_\_\_\_
47. Raw difference favors (i.e., shows more success for) [TX4\_SUCS]
- 1 treatment group
- 2 neither (exactly equal)
- 3 control group
- 4 treatment group, though statistically nonsignificant
- 9 cannot tell

***When significant test information is reported:***

- \_\_\_\_\_
48. t-value (write in the value if available)
- \_\_\_\_\_ a. t-value for reading rate [T\_RATE4]
- \_\_\_\_\_ b. t-value for accuracy [T\_ACC4]
- \_\_\_\_\_ c. t-value for comprehension [T\_COMP4]
- \_\_\_\_\_
49. F-value (write in the value if available)
- \_\_\_\_\_ a. F-value for reading rate [F\_RATE4]
- \_\_\_\_\_ b. F-value for accuracy [F\_ACC4]
- \_\_\_\_\_ c. t-value for comprehension [F\_COMP4]
- \_\_\_\_\_
50. Pretest Effect size for:
- \_\_\_\_\_ a. reading rate [ES4\_PRERATE]
- \_\_\_\_\_ b. reading accuracy [ES4\_PREACC]
- \_\_\_\_\_ c. reading comprehension [ES4\_PRECOMP]
- \_\_\_\_\_
51. Posttest Effect size for:
- \_\_\_\_\_ a. reading rate [ES4\_RATE]
- \_\_\_\_\_ b. reading accuracy [ES4\_ACC]
- \_\_\_\_\_ c. reading comprehension [ES4\_COMP]

Treatment 5

\_\_\_\_\_

52. Sample size [TX5\_N]

\_\_\_\_\_

53. Group pretest mean

\_\_\_\_\_

a. reading rate [PRERATE5]

\_\_\_\_\_

b. accuracy [PREACC5]

c. comprehension [PRECOMP5]

\_\_\_\_\_

54. Treatment group posttest standard deviation

\_\_\_\_\_

a. reading rate [ST5\_PSTRATE]

\_\_\_\_\_

b. reading accuracy [ST5\_PSTACC]

c. comprehension [ST5\_PSTCOMP]

\_\_\_\_\_

55. Group posttest mean

\_\_\_\_\_

a. reading rate [PSTRATE5]

\_\_\_\_\_

b. accuracy [PSTACC5]

c. comprehension [PSTCOMP5]

\_\_\_\_\_

56. Group posttest standard deviation

\_\_\_\_\_

a. reading rate [ST5\_PSTRATE]

\_\_\_\_\_

b. reading accuracy [ST5\_PSTACC]

c. comprehension [ST5\_PSTCOMP]

\_\_\_\_\_

57. Raw difference favors (i.e., shows more success for) [TX5\_SUCS]

1 treatment group

2 neither (exactly equal)

3 control group

4 treatment group, though statistically nonsignificant

9 cannot tell

***When significant test information is reported:***

\_\_\_\_\_

58. t-value (write in the value if available)

\_\_\_\_\_

a. t-value for reading rate [T\_RATE5]

\_\_\_\_\_

b. t-value for accuracy [T\_ACC5]

c. t-value for comprehension [T\_COMP5]

\_\_\_\_\_

59. F-value (write in the value if available)

\_\_\_\_\_

a. F-value for reading rate [F\_RATE5]

\_\_\_\_\_

b. F-value for accuracy [F\_ACC5]

c. t-value for comprehension [F\_COMP5]

\_\_\_\_\_

60. Pretest Effect size for:

\_\_\_\_\_

a. reading rate [ES5\_PRERATE]

\_\_\_\_\_

b. reading accuracy [ES5\_PREACC]

c. reading comprehension [ES5\_PRECOMP]

\_\_\_\_\_

61. Posttest Effect size for:

\_\_\_\_\_

a. reading rate [ES5\_RATE]

\_\_\_\_\_

b. reading accuracy [ES5\_ACC]

c. reading comprehension [ES5\_COMP]

**Treatment 6**

\_\_\_\_\_

62. Sample size [TX6\_N]

- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_
63. Group pretest mean
- a. reading rate [PRERATE6]
  - b. accuracy [PREACC6]
  - c. comprehension [PRECOMP6]
64. Treatment group posttest standard deviation
- a. reading rate [ST6\_PSTRATE]
  - b. reading accuracy [ST6\_PSTACC]
  - c. comprehension [ST6\_PSTCOMP]
65. Group posttest mean
- a. reading rate [PSTRATE6]
  - b. accuracy [PSTACC6]
  - c. comprehension [PSTCOMP6]
66. Group posttest standard deviation
- a. reading rate [ST6\_PSTRATE]
  - b. reading accuracy [ST6\_PSTACC]
  - c. comprehension [ST6\_PSTCOMP]
67. Raw difference favors (i.e., shows more success for) [TX6\_SUCS]
- 1 treatment group
  - 2 neither (exactly equal)
  - 3 control group
  - 4 treatment group, though statistically nonsignificant
  - 9 cannot tell

***When significant test information is reported:***

- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_
68. t-value (write in the value if available)
- a. t-value for reading rate [T\_RATE6]
  - b. t-value for accuracy [T\_ACC6]
  - c. t-value for comprehension [T\_COMP6]
69. F-value (write in the value if available)
- a. F-value for reading rate [F\_RATE6]
  - b. F-value for accuracy [F\_ACC6]
  - c. t-value for comprehension [F\_COMP6]
70. Pretest Effect size for:
- a. reading rate [ES6\_PRERATE]
  - b. reading accuracy [ES6\_PREACC]
  - c. reading comprehension [ES6\_PRECOMP]
71. Posttest Effect size for:
- a. reading rate [ES6\_RATE]
  - b. reading accuracy [ES6\_ACC]
  - c. reading comprehension [ES6\_COMP]

Treatment 7

- \_\_\_\_\_
72. Sample size [TX7\_N]

- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_
73. Group pretest mean
- reading rate [PRERATE7]
  - accuracy [PREACC7]
  - comprehension [PRECOMP7]
74. Treatment group posttest standard deviation
- reading rate [ST7\_PSTRATE]
  - reading accuracy [ST7\_PSTACC]
  - comprehension [ST7\_PSTCOMP]
75. Group posttest mean
- reading rate [PSTRATE7]
  - accuracy [PSTACC7]
  - comprehension [PSTCOMP7]
76. Group posttest standard deviation
- reading rate [ST7\_PSTRATE]
  - reading accuracy [ST7\_PSTACC]
  - comprehension [ST7\_PSTCOMP]
- \_\_\_\_\_
77. Raw difference favors (i.e., shows more success for) [TX7\_SUCS]
- treatment group
  - neither (exactly equal)
  - control group
  - treatment group, though statistically nonsignificant
  - cannot tell

***When significant test information is reported:***

- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_
78. t-value (write in the value if available)
- t-value for reading rate [T\_RATE7]
  - t-value for accuracy [T\_ACC7]
  - t-value for comprehension [T\_COMP7]
79. F-value (write in the value if available)
- F-value for reading rate [F\_RATE7]
  - F-value for accuracy [F\_ACC7]
  - t-value for comprehension [F\_COMP7]
80. Pretest Effect size for:
- reading rate [ES7\_PRERATE]
  - reading accuracy [ES7\_PREACC]
  - reading comprehension [ES7\_PRECOMP]
81. Posttest Effect size for:
- reading rate [ES7\_RATE]
  - reading accuracy [ES7\_ACC]
  - reading comprehension [ES7\_COMP]

**APPENDIX F:  
Excluded Studies**

## APPENDIX F EXCLUDED STUDIES

### 1. ABAB design

- Belfiore, P. J., Skinner, C. H., & Ferkis, M. A. (1995). Effects of response and trial repetition on sight word training for students with learning disabilities. Journal of Applied Behavior Analysis, *28*, 347-348.
- Bolton, A. B. (1991). Reading to read: Evaluating a variant of repeated reading (Doctoral dissertation, University of Southern Mississippi, Hattiesburg, 1991). Dissertation Abstracts International, *52*, 3220A.
- Boyer, L. M. (1991). Effects of reading to read on oral reading fluency. Unpublished Master's project, University of Southern Mississippi.
- Boyer, L. M. (1992). Reading to read: Generalization and effects on comprehension (Doctoral Dissertation, University of Southern Mississippi, Hattiesburg, 1992). Dissertation Abstracts International, *53*, 4945B.
- Cottingham, B. J. (1996). Contributing effects of immediate corrective feedback and previewing in the reading to read intervention. Doctoral dissertation, the University of Southern Mississippi.
- Daly, E. J., Marten, B. K., Dool, E. J., & Eckert, T. L. (1999). A brief experimental analysis for identifying instructional components needed to improve oral reading fluency. Journal of Applied Behavior Analysis, *32*, 83-94.
- Daly, E. J., Marten, B. K., Hamler, K. R., Dool, E. J., & Hintze, J. M. (1998). Using brief functional analysis to select interventions for oral reading. Journal of Behavioral Education, *8*, 203-218.
- Eckert, T. L., Ardoin, S. P., Daisey, D. M., & Scarola, M. D. (2000). Empirically evaluating the effectiveness of reading interventions: The use of brief experimental analysis and single case designs. Psychology in the Schools, *37*, 463-473.
- Kamps, D. M., Barbetta, P. M., Leonard, B. R., & Delquadri, J. D. (1994). Classwide peer tutoring: An integration strategy to improve reading skills and promote peer interactions among students with autism and general education peers. Journal of Applied Behavior Analysis, *27*, 49-61.

- Kastner, J. W. (1997). The utility of reading to read with boys with ADHD during optimal versus nonoptimal methylphenidate effect windows. Doctoral dissertation, University of Southern Mississippi.
- Mefferd, P. E., & Pettegrew, B. S. (1997). Fostering literacy acquisition of students with developmental disabilities: Assisted reading with predictable trade books. Reading Research and Instruction, 36, 177-190.
- Morris, A. L. (1995). An evaluation of listening previewing in the reading to read intervention. Master's thesis. University of Southern Mississippi.
- Prestridge, C. C. (1996). Reading to read and curriculum-based passages: Effects on student performance. Master's thesis, University of Southern Mississippi.
- Swain, K. D., & Allinder, R. M. (1996). The effects of repeated reading on two types of CBM: Computer maze and oral reading with second-grade students with learning disabilities. Diagnostique, 21, 51-66.
- Tingstrom, D. H., Edwards, R. P., & Olmi, D. J. (1995). Listening previewing in reading to read: Relative effects on oral reading fluency. Psychology in the Schools, 32, 318-327.
- Turpie, J. J., & Paratore, J. R. (1995). Using repeated reading to promote reading success in a heterogeneously grouped first grade. In K. A. Hinchman, D. J. Leu, & C. K. Kinzer (Eds.). *Perspectives on Literacy Research and Practice*.
- Walen, S., Hauserman, N., & Dickerson, D. (1974). Comparison of varied reinforcement procedures on reading fluency behavior. Probe: Probing resources of better education/Lida Lee Tall Learning Resources Center, 1, 31-36.
- Weinstein, G., & Cooke, N. L. (1992). The effects of two repeated reading interventions on generalization of fluency. Learning Disability Quarterly, 15, 21-28.

## **2. No fluency measured**

- Bakken, P., Mastropieri, M. A., & Scruggs, T. E. (1997). Reading comprehension of expository science material and students with learning disabilities: A comparison of strategies. The Journal of Special Education, 31, 300-324.
- Brand-Gruwel, S., Aarnoutse, C. A. J., & Van Den Bos, K. P. (1998). Improving text comprehension strategies in reading and listening settings. Learning and Instruction, 8, 63-81.
- Carlton, M. (1981). The effects of intraclass peer tutoring. Doctoral dissertation.



- Dowhower, S. L. & Brown, K. (1992). The effects of predictable materials on first graders' reading comprehension: A teacher action research study. Ohio Reading Teacher, 26, 3-10.
- Garner, C. (1965). Experimental use of the "impress method of reading rehabilitation." ED 003 838.
- Greenwald, M. J., & Kulig, R. (1994). Effects of repeated readings of alphabet books on kindergarteners' letter recognition. Perspectives on Literacy Research and Practice, 25, 231-234.
- Guttentag, R. E. & Haith, M. M. (1980). A longitudinal study of word processing by first grade children. Journal of Educational Psychology, 72, 701-705. (NO fluency and no control)
- Hollingsworth, P. M. (1978). An experimental approach to the impress method of teaching reading. The Reading Teacher, 31, 624-626. (no fluency)
- Kozminsky, L. & Kozminsky, E. (1995). The effects of early phonological awareness training on reading success. Learning and Instruction, 5, 187-201.
- Labbo, L. D., & Teale, W. H. (1990). Cross-age reading: A strategy for helping poor readers. The Reading Teacher, ?, 362-369.
- Lemoine, H. E., Levy, B. A., & Hutchinson, A. (1993). Increasing the naming speed of poor readers: Representations formed across repetitions. Journal of Experimental Child Psychology, 55, 297-328.
- Lorenz, L., & Vockell, E. (1979). Using the neurological impress method with learning disabled readers. Journal of Learning Disabilities, 12, 420-422.
- Lovett, M. W. & Steinback, K. A. (1997). The effectiveness of remedial programs for reading disabled children of different ages: Does the benefit decrease for older children? Learning Disability Quarterly, 20, 189-210.
- O'Shaughnessy, T. E. & Swanson, H. L. (2000). A comparison of two reading intervention for children with reading disabilities. Journal of Learning Disabilities, 33, 257-277.
- Person, M. E., & Burke, D. M. (1984, October). Oral reading in the classroom. Paper presented at the annual meeting of the Northern Rocky Mountain Educational Research Association, Jackson Hole, WY (ERIC Document Reproduction Service ED 251 796).

Peskay, V. E. (1995). Strategy for success: The strengthening of academic skills by means of classroom peer tutoring. Unpublished doctoral dissertation, University of Minnesota.

Reitsma, P. (1988). Reading practice for beginners: Effects of guided reading, reading-while-listening, and independent reading with computer-based speech feedback. *Reading Research Quarterly*, *23*, 219-235.

Spring, C., Blunden, D., & Catheral, M. (1981). Effect on reading comprehension of training to automaticity in word-readings. *Perceptual and Motor Skills*, *53*, 779-786

Sudzina, M., & Foreman, P. (1990). The use of recorded books in a first grade classroom with Chapter I children, *Ohio Reading Teacher*, *25*, 7-14.

### **3. No control group**

Anderson, R. C., Wilkinson, I. A. G., & Mason, J. M. (1991). A microanalysis of the small-group, guided reading lesson: effects of an emphasis on global story meaning. *Reading Research Quarterly*, *26*, 417-441.

Bell, D., M., & Yonker, R. J. (1990). The effects of repeated readings on elementary students' attitude toward reading (self-contained classrooms). (ED 326844).

Bourassa, D. C., Levy, B. A., Dowin, S., & Casey, A. (1998). Transfer effects across contextual and linguistic boundaries: Evidence from poor readers. *Journal of Experimental Child Psychology*, *71*, 45-61.

Dowhower, S. L. (1987). Effect of repeated reading on selected second graders' oral reading and comprehension. Doctoral Dissertation, University of Wisconsin, Madison, WI (Microfile No. 58785DAO, 1986).

Goldstein, A. (1999). Improving students' fluency. Master's Action Research Project, Saint Xavier University and IRI/Skylight. ED 433498.

Hasbrouck, J. E., Ihnot, C., & Rogers, G. H. (1999). "Read Naturally": A strategy to increase oral fluency. *Reading Research and Instruction*, *39*, 27-38.

Herman, P. A. (1985). The effects of repeated readings on reading rate, speech pauses, and word recognition accuracy. *Reading Research Quarterly*, *20*, 553-565.

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*, 94-99.

- Klassen, V. K. (1984). A comparative study of the effects of three methods of oral reading on primary grade pupil reading fluency. Doctoral dissertation, University of Minnesota.
- Knupp, R. (1988). Improving oral reading skills of educationally handicapped elementary school-aged students through repeated readings. Practicum paper, ED 297 275.
- Komisar, K. C. (1999). The effects of direct instruction in prosodic elements on the oral reading fluency of students with learning disabilities. DAI-A 60/04, p.1079.
- LaSalle, A. N. (1999). Bimodal presentation of text, passage previewing, and repeated readings: Effects on reading accuracy, fluency, and comprehension. Doctoral dissertation, Syracuse University.
- Levy, B.A., Abello, B., & Lysynchuk, L. (1997). Transfer from word training to reading in context: Gains in reading fluency and comprehension. Learning Disabilities Quarterly, 20, 173-188.
- Levy, B. A., Nicholls, A., & Kohen, D. (1993). Repeated readings: Process benefits for good and poor readers. Journal of Experimental Child Psychology, 56, 303-327.
- Monda, L. E. (1990). The effects of oral, silent, and listening repetitive reading on the fluency and comprehension of leaning disabled students. DAI-A 50/12-A, p.3919.
- Neville, M. H. (1968). Effects of oral and echoic responses in beginning reading. Journal of Educational Psychology, 59(5), 362-369.
- O'Shea, L. J., Sindelar, P. T., & O'Shea, D. J. (1987). The effects of repeated readings and attentional cues on reading fluency and comprehension of learning disabled readers. Learning Disabilities Research, 2, 103-109.
- 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, 129-142.
- Rasinski, T. V. (1990). Effects of repeated reading and listening while reading on reading fluency. Journal of Educational Research, 83, 147-150.
- Reutzel, D. R., Hollingsworth, P. M., & Eldredge, L. (1994). Oral reading instruction: The impact on student reading development. Reading Research Quarterly, 29, 40-62.

- Roberts, M. L., Turco, T. L., & Shapiro, E. S. (1991). Differential effects of fixed instructional ratios on students' progress in reading. Journal of Psychoeducational Assessment, 1991, 308-318.
- Rokicki, A. L. (1990). The effects of modeling and repeated taped reading instruction on the learning disabled child as ways to improve reading comprehension fluency and word recognition. Practicum paper, ED 335 813.
- Sachs, A. (1984). The effects of previewing activities on oral reading miscue. Remedial and Special Education, 5, 45-49.
- Selvey, A. S. (1990). Effects of Repeated Reading on Decoding Disfluency and Reading Comprehension. DAI-A 51/02, p. 467.
- Sindelar, P. T., Monda, L. E., & O'Shea, L. J. (1990). Effects of repeated readings on instructional- and mastery-level readers. Journal of Educational Research, 83, 220-226.
- Stoddard, K., Valcante, G., Sindelar, P., & O'Shea, L. (1993). Increasing reading rate and comprehension: The effects of repeated readings, sentence segmentation, and intonation training. Reading Research and Instruction, 32(4), 53-65.
- Sutton, P. A. (1991). Strategies to increase oral reading fluency of primary resource students. Practicum paper (ED 335660)
- Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. Journal of Learning Disabilities, 34, 33-58.
- Vaughn, S., Chard, D. J., Bryant, D. P., Coleman, M., Tyler, B., Linan-Thompson, S., & Kouzekanani, K. (2000). Fluency and comprehension interventions for third-grade students. Remdial and Special Education, 21(6), 325-335.
- 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, 59-84.

#### **4. Inadequate quantitative data to compute ES**

- 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. Rowlad (Eds.), *Communications research in learning disabilities and mental retardation* (pp. 33-65). Baltimore, MD: University Park Press.

- Eldredge, J. L. (1990). Increasing the performance of poor readers in the third grade with a group-assisted strategy. Journal of Educational Research, 84, 69-77.
- Faulkner, H. J., & Levy, B. A. (1994). How text difficulty and reader skill interact to produce differential reliance on word and content overlap in reading transfer. Journal of Experimental Child Psychology, 58, 1-24.
- Heckelman, R. (1969). A neurological impress method. Academic Therapy, 4, 277-282.
- Strong, M. W., & Traynelis-Yurek, E. (1983). Behavioral reinforcement within a perceptual-conditioning program of oral reading. ERIC Document Reproduction Service ED 233-328.
- Taylor, N. E., Wade, M. R., & Yekovich, F. R. (1985). The effects of text manipulation and multiple reading strategies on the reading performance of good and poor readers. Reading Research Quarterly, 20, 566-574.
- Van-Bon, W. H. J, Boksebeld, L. M., Font Freide, T. A. M., & van den Hurk, A. J. M. (1991). A comparison of three methods of reading while listening. Journal of Learning Disabilities, 24, 471-476.
- van der Leij, A. (1981). Remediation of reading disabled children by presenting text simultaneously to eye and ear. Bulletin of the Orton Society, 31, 229-243.

## **5. Repeated measures design**

- Bohlen, L. (1988). Building fluency through repeated reading in the third grade. Master's thesis (ED 294158).
- Fleisher, L. S., Jenkins, J. R., & Pany, D. (1979). Effects on poor readers' comprehension of training in rapid decoding. Reading Research Quarterly, 1, 30-46.(Experiment 1)
- Martin, J. G., & Meltzer, R. H. (1976). Visual rhythms: Report on a method for facilitating the teaching of reading. Journal of Reading Behavior, 8, 153-160.
- Rashotte, C. A., & Torgesen, J. K. (1985). Repeated reading and reading fluency in learning disabled children. Reading Research Quarterly, 20, 180-188

## **6. Subjects were not elementary school students.**

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

- Carver, R. P., & Hoffman, J. V. (1981). The effects of practice through repeated reading on gain in reading ability using a computer-based instructional system. *Reading Research Quarterly*, *16*, 374-390.
- Grant, E., & Standing, L. (1989). Effects of rapid decoding training on reading speed and comprehension. *Perceptual and Motor Skills*, *69*, 515-521.
- 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 and Practice*, *15*, 179-189.
- Thomas, A. & Clapp, T. (1989). A comparison of computer-assisted component reading skills training and repeated reading for adolescent poor readers. *Canadian Journal of Special Education*, *5*, 135-144.

## REFERENCES

- Allington, R. L. (1984). Oral reading, In P. D. Pearson (Ed.), *The handbook of reading research* (pp. 829-864). New York: Longman.
- Allington, R. L. (1983a). Fluency: The neglected reading goal in reading instruction. *The Reading Teacher*, 36, 556-561.
- Allington, R. L. (1983b). The reading instruction provided readers of differing reading abilities. *The Elementary School Journal*, 83, 548-559.
- Allington, R. (1977). If they don't read much, how they ever gonna get good? *Journal of Reading*, 21, 57-61.
- Allington, R.L. Stuetzel, H., Shake, M., & Lamarche, S. (1986). What is remedial reading? A descriptive study. *Reading Research and Instruction*, 24, 15-30.
- \*Allor, J. H., Fuchs, D., & Mathes, P. G. (2001). Do students with and without lexical retrieval weaknesses respond differently to instruction? *Journal of Learning Disabilities*, 34(3), 264-275.
- Allington, R. L., & McGill-Frazen, A. (1989). School response to reading failure: Instructon for Chapter 1 and special education students in grades two, four, and eight. *The Elementary School Journal*, 89, 530-542.
- Anderson, B. (1981). The missing ingredient: Fluent oral reading. *The Elementary School Journal*, 81(3), 173-177.
- Anderson, R.C., Hiebert, E. H., Scott, J. A., & Wilkerson, I. A. (1985). *Becoming a nation of readers: The report of the Commission on Reading*. Champaign, IL: The Center for the Study of Reading.
- Anderson, R. C., Wilkinson, I. A. G., & Mason, J. M. (1991). A microanalysis of the small-group, guided reading lesson: effects of an emphasis on global story meaning. *Reading Research Quarterly*, 26, 417-441.
- \*Arnold, R. D. (1972). A comparison of the neurological impress method, the language experience approach, and classroom teaching for children with reading disabilities. Final report, Washington DC: National Center for Educational Research and Development (ED 073428.)
- \*Aslett, R. (1990). Effects of the oral recitation lesson on reading comprehension of fourth grade developmental readers. Unpublished doctoral dissertation, Brigham Young University, Provo, UT.

- \*Baker, S., Gersten, R., & Keating, T. (2000). When less may be more: A 2-year longitudinal evaluation of a volunteer tutoring program requiring minimal training. Reading Research Quarterly, 35(4), 494-519.
- Biemiller, A. (1977-1978). Relationship between oral reading rates for letters, words, and simple text in the development of reading achievement. Reading Research Quarterly, 13, 223-253.
- Blanchard, J. S. (1980). Preliminary investigation of transfer between single-word decoding ability and contextual reading comprehension by poor readers in grade six. Perceptual and Motor Skills, 51, 1271-1281.
- Bohlen, L. (1988). Building fluency through repeated reading in the third grade. Master's thesis (ED 294158).
- Boder, E., & Jarrico, S. (1982). The Boder Test of Reading-Spelling Patterns. PsychCorp, A brand of Harcourt Assessment, Inc., San Antonio, TX.
- Bolton, A. B. (1991). Reading to read: Evaluating a variant of repeated reading (Doctoral dissertation, University of Southern Mississippi, Hattiesburg, 1991). Dissertation Abstracts International, 52, 3220A.
- Bos, C. S. (1982). Getting past decoding: Assisted and repeated readings as remedial methods for learning disabled students. Topics in Learning and Learning Disabilities, 1, 51-57.
- Boyer, L. M. (1991). Effects of reading to read on oral reading fluency. Unpublished Master's project, University of Southern Mississippi.
- Boyer, L. M. (1992). Reading to read: Generalization and effects on comprehension (Doctoral Dissertation, University of Southern Mississippi, Hattiesburg, 1992). Dissertation Abstracts International, 53, 4945B.
- Bristow, P. S. (1985). Are poor readers passive readers? Some evidence, possible explanations and potential solutions. The Reading Teacher, 39(3), 318-324.
- Carbo, M. (1978). Teaching reading with talking books. The Reading Teacher, 32, 267-273.
- Carver, R. P., & Hoffman, J. V. (1981). The effects of practice through repeated reading on gain in reading ability using a computer-based instructional system. Reading Research Quarterly, 16, 374-390.
- Chomsky, C. (1978). When you still can't read in third grade: After decoding, what? In S. J. Samuels (Ed.), What research has to say about reading instruction (pp. 13-30). Newark, DE: International Reading Association.



- \*Cohen, A. L. (1989). An Evaluation of the Effectiveness of Two Methods for Providing Computer-Assisted Repeated Reading Training to Reading Disabled Students. DAI-A 49/07. P. 1746.
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2<sup>nd</sup> Ed.). Hillsdale, NH: Erlbaum.
- \*Collins, V. L. (1994). Improving children's reading skills using previewing, repeated readings, and self-monitoring techniques. Unpublished doctoral dissertation.
- \*Conte, R., & Humphreys, R. (1989). Repeated reading: Using audio-taped material enhances oral reading in children with reading difficulties. Journal of Communication Disorders, 22, 65-79.
- Cooper, H., & Hedges, L. V. (1994). The handbook of research synthesis. New York: Russell Sage Foundation.
- Cordray, D. S., & Sonnefeld, L. J. (1985). Quantitative synthesis: An actuarial base for planning research. New Directions in Program Evaluation, 27, 29-48.
- Cottingham, B. J. (1996). Contributing effects of immediate corrective feedback and previewing in the reading to read intervention. Doctoral dissertation, the University of Southern Mississippi.
- Cromer, W. (1970). The difference model: A new explanation for some reading difficulties. Journal of Educational Psychology, 61, 471-483.
- 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. Rowlad (Eds.), *Communications research in learning disabilities and mental retardation* (pp. 33-65). Baltimore, MD: University Park Press.
- \*Denton, C. A. (1997). An evaluation of an implementation of the Reading Recovery Program. Paper presented at the Annual Meeting of the Southwest Educational Research Association, Austin, TX. (ED 408 309)
- Donahue, P. L., Finnegan, R. J., Lutkus, A. D., Allen, N. L., & Campbell, J. R. (2001). The nation's report card: Fourth-grade reading 2000. Executive summary. [WWW document]. URL <http://nces.ed.gov/nationsreportcard/pubs/main2000/2001499.asp>
- Dowhower, S. L. (1987). Effect of repeated reading on selected second graders' oral reading and comprehension. Doctoral Dissertation, University of Wisconsin, Madison, WI (Microfile No. 58785DAO, 1986).

- Dowhower, S. L. (1991). Speaking of prosody: Fluency's unattended bedfellow. Theory in Practice, 30(3), 158-164.
- Durkin, D. (1993). Teaching them to read (6th Ed.). Needham, MA: Allyn and Bacon.
- Ekwall Reading Inventory, Ekwall, E. (1979). Needham Heights, MA: Allyn and Bacon.
- \*Eldredge, J. L., Reutzel, D. R., & Hollingsworth, P. M. (1996). Comparing the effectiveness of two oral reading practices: round-robin reading and the shared book experience. Journal of Literacy Research, 28(2), 442-453.
- Embrey, E. (1968). A study on the effectiveness of the neurological impress method as a remedial reading technique. Unpublished master's thesis, Central Washington State College.
- Faulkner, H. J., & Levy, B. A. (1999). Fluent and nonfluent forms of transfer in reading: Words and their message. Psychonomic Bulletin and Review, 6, 111-116.
- \*Fleisher, L. S., Jenkins, J. R., & Pany, D. (1979). Effects on poor readers' comprehension of training in rapid decoding. Reading Research Quarterly, 1, 30-46.
- 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, D., Fuchs, L. S., Mathes, P., Lipsey, M., & Roberts, H. (2001). Does LD exist? Learning Disabilities Summit--Building a Foundation for the Future. Sponsored by the Office of Special Education Programs, U.S. Department of Education, Washington, D.C.
- Fuchs, D., & Fuchs, L.S. (1998). Researchers and teachers working together to adapt instruction for diverse learners. Learning Disabilities Research and Practice, 13, 126-137.
- Fuchs, D., Fuchs, L.S., & Burish, P. (2000). Peer-assisted learning strategies: An evidence-based practice to promote reading achievement. Learning Disabilities Research and Practice, 15(2), 85-91.
- \*Fuchs, D., Fuchs, L.S., Mathes, P.G., & Simmons, D.C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. American Educational Research Journal, 34(1), 174-206.

- \*Fuchs, D., Fuchs, L. S., Thompson, A., Others. 1998-99 first grade PALS (PALS, PALS+Fluency, PALS+Comprehension, & Control)
- \*Fuchs, D., Fuchs, L.S., & Others. 1999-2000 first grade PALS study (PALS, PALS + Fluency, & Control).
- \*Fuchs, D., Fuchs, L.S., & Others. 2000-2001 first grade PALS study (PALS, PALS + Fluency, & Control).
- \*Fuchs, D., Fuchs, L.S., & Others. 1999-2000 kindergarten PALS study (Decoding PALS, Sound Play, Decoding PALS + Sound Play, & Control)
- Fuchs, D., Fuchs, L.S., Thompson, A., Al Otaiba, S., Yen, L., Yang, N.J., Braun, M., & O'Connor, R.E. (2001). Is reading important in reading-readiness programs: A randomized field trial with teachers as program implementers. Journal of Educational Psychology, 93(2), 251-267.
- Fuchs, D., Fuchs, L.S., Thompson, A., Svenson, E., Yen, L., Al Otaiba, S., Yang, N., Nyman, K., Prentice, K., Kazdan, S., & Saenz, L. (2001). Peer-assisted learning strategies in reading: Extensions to kindergarten, first grade, and high school. Remedial and Special Education, 22(1), 15-21.
- Gambrell, L. B., Wilson, R. M., & Gantt, W. N. (1981). Classroom observations of task-attending behaviors of good and poor readers. Journal of Educational Research, 74, 400-404.
- Gleser, L. J., & Olkin, I. (1994). Stochastically dependent effect sizes. In Cooper, H., & Hedges, L. V. (Eds.). The Handbook of Research Synthesis. The Russell Sage Foundation.
- Goldstein, A. (1999). Improving students' fluency. Master's Action Research Project, Saint Xavier University and IRI/Skylight. ED 433498.
- Golinkoff, R. M., & Rosinski, R. R. (1976). Decoding, semantic processing, and reading comprehension skill. Child Development, 47, 252-258.
- Golinkoff, R. (1975-76). Comprehension in good and poor readers. Reading Research Quarterly, 4, 623-656.
- Good, R. H., Simmons, D. C., & Smith, S. (1998). Effective academic interventions in the United States: Evaluating and enhancing the acquisition of early reading skills. School Psychology Review, 27, 45-56.
- Grant, E., & Standing, L. (1989). Effects of rapid decoding training on reading speed and comprehension. Perceptual and Motor Skills, 69, 515-521.

- Greenwood, C. R., Delquadri, J. C., & Carta, J. J. (1988). Classwide peer tutoring. Delray Beach, FL: Educational Achievement System.
- \*Gunn, B., Biglan, A., Smolkowski, K., & Ary, D. (2000). The efficacy of supplemental instructional in decoding skills for Hispanic and non-Hispanic students in early elementary school. The Journal of Special Education, 34(2), 90-103.
- Guttentag, R. E., & Haith, M. M. (1978). Automatic processing as a function of age and reading ability. Child Development, 49, 707-716.
- \*Hannah, A. M. (1994). The effects of repeated readings on the reading fluency and comprehension of second-grade, Chapter I students. Unpublished Ph.D dissertation, University of Delaware.
- Harris, T. L., & Hodges, R. E. (1981). A Dictionary of Reading and Related Terms. Newark, NJ: International Reading Association.
- Haynes, M. C., & Jenkins, J. R. (1986). Reading instruction in special education resource rooms. American Educational Research Journal, 23, 161-190.
- Heckelman, R. (1969). A neurological impress method. Academic Therapy, 4, 277-282.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. Journal of Educational Statistics, 6, 107-128.
- Hedges, L. V. (1994). Fixed effects models. In H. Cooper & L. V. Hedges (Eds.), The handbook of research synthesis (pp. 285-299). New York: Russell Sage Foundation.
- Hedges, L.V., & Olkin, I. (1985). Statistical methods for meta-analysis. Orlando, FL: Academic Press.
- Herman, P. A. (1985). The effects of repeated readings on reading rate, speech pauses, and word recognition accuracy. Reading Research Quarterly, 20, 553-565.
- Hoffman, J. R. (1987). Rethinking the role of oral reading in basal instruction. The Elementary School Journal, 87, 367-373.
- Hoffman, J. R., Isaacs, M. E., Roser, N., & Farest, C. (1989, November-December). Developing reading fluency and story comprehension using a recitation lesson format. Paper presented at the 39<sup>th</sup> Annual Meeting of the National Reading Conference, Austin, TX.

- Hollingsworth, P. M. (1978). An experimental approach to the impress method of teaching reading. The Reading Teacher, 31, 624-626.
- \*Hollingsworth, P. (1970). An experiment with the impress method of teaching reading. The Reading Teacher, 24, 112-114.
- 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, 94-99.
- \*Hunt, G. (1994). The improvement of reading comprehension of the second grade at-risk students using multisensory methods of instruction. M.A. project, Saint Xavier University. (ED 371 293)
- Isakson, R. L., & Miller, J. W. (1976). Sensitivity to syntactic and semantic cues in good and poor comprehenders. Journal of Educational Psychology, 68, 787-792.
- \*Jones, K. M., Torgesen, J. K., & Sexton, M. A. (1987). Using computer guided practice to increase decoding fluency in learning disabled children: A study using the Hint and Hunt I Program. Journal of Learning Disabilities, 20, 122-128.
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grade. Journal of Educational Psychology, 80, 437-447.
- Katz, L., & Wicklund, D. A. (1971). Word scanning rate for good and poor readers. Journal of Educational Psychology, 62, 1380-140.
- \*King, R. T. (1982). Learning from PAL. Reading Teacher, 35(6), 682-685.
- \*Koch, K. A. (1984). The effects of oral repeated reading on second grade students' oral reading accuracy, rate, and comprehension. Unpublished Ph.D. dissertation, Purdue University.
- \*Koskinen, P. A., & Blum, I. H. (1984). Repeated oral reading and the acquisition of fluency. In J. A. Niles & L. A. Harris (Eds.), *Changing perspectives on research in reading/language processing and instruction: Thirty-third yearbook of the National Reading Conference* (pp. 183-187). Rochester, NY: National Reading Conference.
- Koskinen, P. A., & Blum, I. H. (1986). Paired repeated reading: A classroom strategy for developing fluent reading. The Reading Teacher, 40, 70-75.

- Komisar, K. C. (1999). The effects of direct instruction in prosodic elements on the oral reading fluency of students with learning disabilities. DAI-A 60/04, p.1079.
- Kuhn, M. R., & Stahl, S. A. (2000). Fluency: A review of developmental and remedial practices. CIERA Report #2-008.
- LaBerge, D., & Samuel, S. J. (1974). Toward a theory of automatic information processing in reading. Cognitive Psychology, 2, 293-323.
- \*Laffey, J. L., Kelly, D., & Perry, B. (1979). The effects of repeated reading of taped literature on reading achievement. Paper presented at the Annual Meeting of the National Reading Conference, San Antonio, TX.
- Langford, K., Slade, K., & Barnett, A. (1974). An examination of the impress technique in remedial reading. Academic Therapy, 9, 303-309.
- Leu, D. J., DeGross, L. C., & Simmons, H. D. (1986). Predictable texts and interactive-compensatory hypotheses: Evaluating individual differences in reading ability, context use and comprehension. Journal of Educational Psychology, 78, 347-352.
- Levy, B. A., Nicholls, A., & Kohen, D. (1993). Repeated readings: Process benefits for good and poor readers. Journal of Experimental Child Psychology, 56, 303-327.
- Lipsey, M. W., & Wilson, D. B. (2000). Practical meta-analysis – Applied social research methods. Thousand Oaks, CA: Sage Publications, Inc.
- Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. Reading and Writing Quarterly: Overcoming Learning Disabilities, 13, 123-146.
- Logan, G. D. (1988). Toward an instance theory of automatization. Psychological Review, 95, 492-527.
- Logan, G. D. (1985). Skill and automaticity: Relations, implications, and future directions. Canadian Journal of Psychology, 39, 367-386.
- Lorenz, L., & Vockell, E. (1979). Using the neurological impress method with learning disabled readers. Journal of Learning Disabilities, 12, 420-422.
- \*Marston, D., Deno, S. L., & Kim, D., Diment, K., & Rogers, D. (1995). Comparison of reading intervention approaches for students with disabilities. Exceptional Children, 62, 20-37.

- \*Mathes, P. G., & Babyak, A. E. (2001). The effects of peer-assisted literacy strategies for first-grade readers with and without additional mini-skills lessons. Learning Disabilities Research and Practice, 16(1), 28-44.
- Mastropieri, M. A., Linart, A., & Sruggs, T. (1999). Strategies to increase reading fluency. Intervention in School and Clinic, 34(5), 278-282.
- \*Mathes, P. G., Howard, J. K., Allen, S. H., & Fuchs, D. (1998). Peer-assisted learning strategies for first graders: Responding to the needs of diverse learners. Reading Research Quarterly, 33(1), 62-94.
- \*Mathes, P. G., & Fuchs, L. S. (1993). Peer-mediated reading instruction in special education resource rooms. Learning Disabilities Research and Practice, 8, 233-243 (ED 408804).
- Mathes, P. G., Simmons, D. C., & Davis, B. I. (1992). Assisted reading techniques for developing reading fluency. Reading Research and Instruction, 31(4), 70-77.
- McGill-Frazen, A., & Allington, R. (1991). The gridlock of low reading achievement: Perspectives on practice and policy. Remedial & Special Education, 12, 20-30.
- \*Miller, A., Robson, D., & Bushell, R. (1986). Parental participation in paired reading: A controlled study. Educational Psychology, 6, 277-284.
- Miller, K., & Snow, D. (2004). Noteworthy perspectives: Out-of-school time programs for at-risk students. Aurora, CO: Mid-continent Research for Education and Learning.
- Morris, A. L. (1995). An evaluation of listening previewing in the reading to read intervention. Master's thesis. University of Southern Mississippi.
- Moyer, S. B. (1982). Repeated reading. Journal of Learning Disabilities, 15, 619-623.
- National Reading Panel (2000). Report of the subgroups: National Reading Panel. Washington, DC: National Institute of Child Health and Development.
- Nelson, L., & Morris, D. (1986). Supported oral reading: A year-long intervention study in two inner-city primary grade classrooms. Paper presented at the annual meeting of the National Reading Conference, Austin, TX.
- Oaken, R., Wiener, M., & Cromer, W. (1978). Identification, organization, and reading comprehension for good and poor readers. Journal of Educational Psychology, 62, 71-78.

- O'Shea, L. J., Sindelar, P. T., & O'Shea, D. J. (1987). The effects of repeated readings and attentional cues on reading fluency and comprehension of learning disabled readers. Learning Disabilities Research, 2, 103-109.
- 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, 129-142.
- Pace, A. J., & Golinkoff, R. M. (1976). The relationship between word difficulty and access of single-word meaning by skilled and less skilled readers. Journal of Educational Psychology, 68, 760-767.
- Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skills. Journal of Educational Psychology, 67, 461-469.
- 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: Office of Educational Research and Improvement, U.S. Department of Education.
- Prestridge, C. C. (1996). Reading to read and curriculum-based passages: Effects on student performance. Master's thesis, University of Southern Mississippi.
- \*Rashotte, C. A., MacPhee, K., & Torgesen, J. K. (2001). The effectiveness of a group reading instruction program with poor readers in multiple grades. Learning Disability Quarterly, 24, 119-134.
- Rashotte, C. A., & Torgesen, J. K. (1985). Repeated reading and reading fluency in learning disabled children. Reading Research Quarterly, 20, 180-188.
- Rasinski, T. V. (1989). Fluency for everyone: Incorporating fluency instruction in the classroom. The Reading Teacher, May, 690-693.
- Rasinski, T. V. (1990). Effects of repeated reading and listening while reading on reading fluency. Journal of Educational Research, 83, 147-150.
- Rasinski, T. V. (1990a). Effects of repeated reading and listening while reading on reading fluency. Journal of Educational Research, 83, 147-150.
- \*Rasinski, T. V., Padak, N., Linek, W., & Sturtevant, B. (1994). Effects of fluency development on urban second-grade readers. Journal of Educational Research, 87, 158-165.
- Raudenbush, S. W. (1994). Random effects models. In H. Cooper & L. V. Hedges (Eds.), The handbook of research synthesis (pp. 301-321). New York: Russell Sage Foundation



- Reitsma, P. (1988). Reading practice for beginners: Effects of guided reading, reading-while-listening, and independent reading with computer-based speech feedback. Reading Research Quarterly, *23*, 219-235.
- \*Reutzel, D. R., & Hollingsworth, P. M. (1993). Effects of fluency training on second graders' reading comprehension, Journal of Educational Research, *86*(6), 325-331.
- \*Reutzel, D. R., Hollingsworth, P. M., & Eldredge, L. (1994). Oral reading instruction: The impact on student reading development. Reading Research Quarterly, *29*, 40-62.
- \*Richek, M. A., & McTeague, B. K. (1988). The "Curious George" strategy for students with reading problems. The Reading Teacher, *42*, 220-226.
- Rose, T. L. (1984a). The effects of two prepractice procedures on oral reading. Journal of Learning Disabilities, *17*, 544-548.
- Rose, T. L. (1984b). The Effects of previewing on retarded learners' oral reading. Education and Training of the Mentally Retarded, *19*(1), 49-53
- Rose, T. L. (1984c). Effects of previewing on the oral reading of mainstreamed behaviorally disordered students. Behavioral Disorders, *10*(1), 33-39
- Rose, T. L., & Beattie, J. R. (1986). Relative effects of teacher-directed and taped previewing on oral reading. Learning Disabilities Quarterly, *9*, 193-199.
- Rose, T. L., & Sherry, L. (1984). Relative effects of two previewing procedures on LD adolescents' oral reading performance. Learning Disabilities Quarterly, *7*, 39-44.
- Royer, J. M. (1997). A cognitive perspective on the assessment, diagnosis, and remediation of reading skills. In G. Phye (Ed.), *Handbook of academic learning* (pp. 199-234). San Diego, CA: Academic Press.
- Royer, J. M., & Sinatra, G. M. (1994). A cognitive theoretical approach to reading diagnostics. Educational Psychology Review, *6*, 81-113.
- Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), Attention and performance (Vol. 6, pp. 573-603). New York: Academic Press.
- Salend, S. J., & Nowak, M. R. (1988). Effects of peer previewing on LD students oral reading skills. Learning Disabilities Quarterly, *11*, 47-53.
- Samuel, S. J. (1979). The method of repeated readings. The Reading Teacher, *32*, 403-408.

- Schreiber, P. A. (1980). On the acquisition of reading fluency. Journal of Reading Behavior, 12, 177-186.
- Schreiber, P. A. (1987). Prosody and structure in children's syntactic processing. In R. Horowitz & S. J. Samuels (Eds.), Comprehending oral and written language. New York: Academic Press.
- Scruggs, T. E., & Mastropieri, M. A. (1998). Tutoring students with special needs. In K. Topping & S. Ehly (Eds.), Peer assisted learning (pp. 165-182). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Selvey, A.S. (1990). Effects of repeated reading on decoding disfluency and reading comprehension. DAI-A 51/02, p. 467.
- \*Shany, M. T., & Biemiller, A. (1995). Assisted reading practice: Effects on performance for readers in grades 3 and 4. Reading Research Quarterly, 30, 382-395.
- \*Simmons, D. C., Fuchs, L. S., Fuchs, D., Mathes, P. G., & Hodge, J. P. (1995). Effects of explicit teaching and peer tutoring on the reading achievement of learning disabled and low-performing students in regular classrooms. Elementary School Journal, 95(5), 387-408.
- \*Simmons, D. C., Fuchs, D., Fuchs, L. S., Hodge, J. P., & Mathes, P. G. (1994). Importance of instructional complexity and role reciprocity to classwide peer tutoring. Learning Disabilities Research & Practice, 9, 203-212.
- Sindelar, P.T.(1987). Increasing reading fluency. Teaching Exceptional Children, 19, 59-60.
- \*Sindelar, P. T. (1982). The effects of cross-aged tutoring on the comprehension skills of remedial reading students. Journal of Special Education, 16(2), 199-206.
- Sindelar, P. T., Monda, L. E., & O'Shea, L. J. (1990). Effects of repeated readings on instructional- and mastery-level readers. Journal of Educational Research, 83, 220-226.
- Smith, F., & Holmes, D. L. (1970-1971). The independence of letter, word and meaning identification in reading. Reading Research Quarterly, 6, 394-415.
- Snow, C. E., Burns, S., & Griffin, P. (Eds.). (1998). Preventing reading difficulties in young children. Washington, DC: National Academy Press.

- Spear-Swerling, L., & Sternber, R. J. (1994). The road not taken: An integrativetheoretical model of reading disability. Journal of Learning Disabilities, 27, 91-103.
- Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. Reading Research Quarterly, 16, 32-71.
- Stanovich, K. E., & Stanovich, P. J. (1995). How research might inform the debate about early reading acquisition. Journal of Research in Reading, 18, 87-105.
- Stanovich, K. E. (1990). Concepts in developmental theories of reading skill: Cognitive resources, automaticity, and modularity. Developmental Review, 10, 72-100.
- Stanovich, K. E., & West, R. F. (1981). The effects of sentence context on Ongoing word recognition: Tests of a two-process theory. Journal of Experimental Psychology: Human Perception and Performance, 7, 658-672.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. Reading Research Quarterly, 21, 360-407.
- Steiner, R., Wiener, M., & Cromer, W. (1971). Comprehension training and Identification for good and poor readers. Journal of Educational Psychology, 62, 506-513.
- Stoddard, K., Valcante, G., Sindelar, P., & O'Shea, L. (1993). Increasing reading rate and comprehension: The effects of repeated readings, sentence segmentation, and intonation training. Reading Research and Instruction, 32(4), 53-65.
- Swain, K. D., & Allinder, R. M. (1996). The effects of repeated reading on two types of CBM: Computer maze and oral reading with second-grade students with learning disabilities. Diagnostique, 21, 51-66.
- Tavegia, T. C. (1974). Resolving research controversy through empirical cumulation: Toward reliable sociological knowledge. Sociological Methods and Research, 2, 395-407.
- The Commission on Chapter 1. "Making Schools Work for Children in Poverty." Education Week 12, 16 (January 13, 1993): 46-48.
- Tingstrom, D. H., Edwards, R. P., & Olmi, D. J. (1995). Listening previewing in reading to read: Relative effects on oral reading fluency. Psychology in the Schools, 32, 318-327.

- Topping, K. (1989). Paired learning: Tutoring by non-teachers. The Paired Reading Bulletin, No. 5. Available from the Oastler Centre, 103 New St. Huddersfield, HD12VA, West Yorkshire, England.
- Turpie, J. J., & Paratore, J. R. (1995). Using repeated reading to promote reading success in a heterogeneously grouped first grade. In K. A. Hinchman, D. J. Leu, & C. K Kinzer (Eds.). *Perspectives on Literacy Research and Practice*.
- Twenty-second Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act (IDEA) (2000). US Department of Education, Washington, DC.
- Vaughn, S., Chard, D. J., Bryant, D. P., Coleman, M., Tyler, B., Linan-Thompson, S., & Kouzekanani, K. (2000). Fluency and comprehension interventions for third-grade students. Remdial and Special Education, 21(6), 325-335.
- \*Vollands, S. R., Topping, K. J., & Evans, R. M. (1999). Computerized self-assessment of reading comprehension with the accelerated reader: Action research. Reading & Writing Quarterly, 15, 197-211
- Weber, G. (1971). Inner city children can be taught to read: Four successful schools (CGE Occasional Papers No. 18). Washington, DC: Council for Basic Education. (ERIC Document Reproduction Service No. Ed 057 125)
- Weinstein, G., & Cooke, N. L. (1992). The effects of two repeated reading interventions on generalization of fluency. Learning Disability Quarterly, 15, 21-28.
- Wilson, S. J. (2000). Effectiveness of school violence prevention programs: Dissertation, Vanderbilt University.
- Woodcock, R.W. (1973). Woodcock reading mastery tests. Circle Pines, Minnesota: American Guidance Service.
- Yang, J. (2000). A meta-analytical review of the effects of repeated reading with remedial readers and students with disabilities. Vanderbilt University.
- Ysseldyke, J. E., O'Sullivan, P. S., Thurlow, M. L., & Christensen, S. L. (1989). Are different kinds of instructional tasks used by different categories of students in different settings? School Psychology Review, 18(1), 98-111
- Zbrodoff, N. J., & Logan, G. D. (1986). On the autonomy of mental processes: A case study of arithmetic. Journal of Experimental Psychology: General, 115, 118-130.