8-2021

A Multilevel Meta-Analysis of Paired Oral Reading Methods in Elementary Classrooms

Jacob D. Downs
Utah State University

A MULTILEVEL META-ANALYSIS OF PAIRED ORAL READING METHODS IN ELEMENTARY CLASSROOMS

by

Jacob D. Downs

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Education

Approved:

Kathleen A. J. Mohr, Ed.D. Major Professor	Cindy D. Jones, Ph.D. Committee Member
Parker Fawson, Ed.D. Committee Member	Nicole Pyle, Ph.D. Committee Member
Tyson S. Barrett, Ph.D. Committee Member	Chase Young, Ph.D. Committee Member

D. Richard Cutler, Ph.D.
Interim Vice Provost of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2021

ABSTRACT

A Multilevel Meta-analysis of Paired Oral Reading Methods in Elementary Classrooms

by

Jacob D. Downs, Doctor of Philosophy
Utah State University, 2021

Major Professor: Kathleen A. J. Mohr, Ed.D.
Department: School of Teacher Education and Leadership

Popularity among various methods of Paired Oral Reading has waxed and waned during recent decades. Ostensibly, these methods share a common trait—a higher level tutor reading aloud synchronously with a lower-level tutee; however, previous research has not attempted to synthesize across methods. This study systematically reviewed and meta-analyzed four methods of Paired Oral Reading; Neurological Impress Method, Paired Reading, Dyad Reading, and Read Two Impress. A systematic search was conducted across five academic databases to identify studies reporting tutee outcomes from Paired Oral Reading. Studies meeting the inclusion criteria were systematically coded and effect sizes were calculated for tutee fluency and comprehension outcomes. Results from the multivariate, multilevel meta-analysis with meta-regression indicated that Paired Oral Reading methods are effective as supporting tutee reading outcomes (g = 0.58). Results also include higher outcomes for comprehension (g = 0.65) than fluency (g = 0.48) and higher outcomes from adult tutors (g = 0.73) than cross-age (g = 0.26) or peer

tutors (g = 0.52). Results further indicated duration of intervention in weeks (b = 0.04) as a significant moderator of Paired Oral Reading Outcomes. Implications include future design recommendations for meta-analytic literacy research and considerations for classroom implementation. Further, synthesizing these four related methods into a new, single field of research is proposed: Synchronous Partnered Oral Reading Techniques (SPORT).

(173 pages)

PUBLIC ABSTRACT

A Multilevel Meta-analysis of Paired Oral Reading Methods in Elementary Classrooms

Jacob D. Downs

For decades, researchers and practitioners have supported developing readers via synchronous oral reader with a stronger peer or adult tutor. These methods–collectively known as Paired Oral Reading—are purported to promote reading achievement for tutee readers. However, despite nearly 60 years of investigation, no known research has adequately reviewed and synthesized the effects of similar practices commonly known as Paired Oral Reading. This dissertation systematically reviewed the published literature on the various methods of Paired Oral Reading with elementary-age students then meta-analyzed the quantitative studies that met pre-established inclusion criteria. The meta-analysis calculated the effect of Paired Oral Reading outcomes on tutee reading achievement and subsequently investigated differential outcomes for various factors such as tutor type, time variables, and reading outcome (fluency/comprehension). The results indicate that tutees receiving Paired Oral Reading support experience greater achievement with adult tutors over peer tutors and in reading comprehension versus oral reading fluency. The systematic review and meta-analysis of these data demonstrate that strategic use of these methods can support weak or developing elementary readers. Further, it is proposed that future research and application of these techniques are framed as a single family of methods; Synchronous Partnered Oral Reading Techniques

(SPORT). Other implications for future research and classroom application of SPORT are discussed.

DEDICATION

Dedicated to Emily, my wife and partner. She earned a doctorate of her own throughout my graduate school experience.

ACKNOWLEDGMENTS

Kit, at my initial orientation you said doc students were in charge of "driving their own bus." I hope this is what you had mind.

To my advisor, Dr. Mohr, I am grateful for your mentorship. I spent countless hours in your office asking question after question. You listened, responded, and always set aside your work to help me with mine. You graciously reviewed multiple drafts of the various manuscripts I wrote as a doctoral student, even (perhaps especially) the poorly written attempts. You lifted me to a higher plane of thought by challenging me throughout the past four years. I cannot repay you for the significant time I spent in your tutelage. Thank you.

To my committee, thank you for your leadership and mentorship throughout my doctoral experience. I learned unique and important lessons for each of you. Dr. Jones, you introduced to me literacy research and all its possibilities. I distinctly remember first hearing the term 'Hedge's *g*' in your class. Dr. Fawson, you introduced me literacy leadership, systems-level thinking, and challenged me to read NRP 2000 in its entirety. Dr. Pyle, you taught me how to design a research study and inspired what would later become my first publication, Downs et al., (2020). Dr. Barrett, you taught me multiple regression, inspired my journey into multilevel modeling, and introduced me to the wonderland of R. Dr. Young, you taught me the importance of bridging research into practice and helped establish my effort in that realm, *Teaching Literacy Podcast.* Thank you all.

To my colleagues at Cache County School District, too numerable to mention,

thank you for your camaraderie and leadership. Teaching young minds to read is among the most noble endeavors. Your unyielding emphasis to this cause is an inspiration to schools everywhere. White Pine Faculty, Instructional Coaches, and now the K-6 Literacy Team, you are all family.

To my family, thank you for your support throughout this process. Emily, thank you for your sacrificing your time for me to complete this degree. I look forward to being home for dinner. Nolan and Elliot, thanks for understanding that Dad needed to be gone a lot to work on school. To Mom, Dad, Bill, and Nadine, thank you for being cheerleaders and asking about my paper.

Last, thank you to those scholars who challenged my thinking and continue to inspire me: Walter Kintsch, D. Ray Reutzel, Elfrieda Hiebert, P. David Pearson, Lloyd Eldredge, Nell K. Duke, Timothy Rasinski, Tim Shanahan, Michael Pressley, Gerald Duffy, J. Joop Hox, and, of course, L.S. Vygotsky.

Jacob D. Downs

CONTENTS

Page

ABSTRACT iii

PUBLIC ABSTRACT v

DEDICATION vii

ACKNOWLEDGMENTS viii

LIST OF TABLES xii

LIST OF FIGURES xiii

CHAPTER 1: INTRODUCTION 1

- Dysfluent Readers 2
- Tutored Support in Reading Fluency 3
- Problem Statement 5
- Purpose of the Meta-Analytic Study 7
- Conclusion 10
- Definition of Terms 11

CHAPTER 2: LITERATURE REVIEW 13

- Search Process 13
- Categorizing Paired Oral Reading Techniques 14
- History and Evolution of Paired Oral Reading Methods 15
- Common Features of Paired Oral Reading 26
- Reported Academic Outcomes of Paired Oral Reading 31
- Outcomes in Tutee Reading Attitudes 46
- Limitations of the Paired Oral Reading Literature 49
- Conclusion 52

CHAPTER 3: METHODS 54

- Research Questions and Method 54
- Data Analysis 59
- Conclusion 68

Page

CHAPTER 4: RESULTS 70

Study Coding 70
Description of the Study Sample 71
Assessing Initial Model Fit 85
Results by Research Question 89
Limitations 96
Summary of Findings 99

CHAPTER 5: DISCUSSION AND CONCLUSION 100

Summary of Multilevel Outcomes 100
Discussion of Multilevel Outcomes 101
Effects of Synchronous Partnered Oral Reading Techniques 102
Implications for Research and Instruction 108
Conclusion 127

REFERENCES 129

APPENDIX: CODING MANUAL 147

CURRICULUM VITAE 153

LIST OF TABLES

		Page
Table 1.1	Description of Paired Oral Reading Methods	4
Table 3.1	Description of Moderator Analysis Variables	66
Table 4.1	Descriptive Statistics for Reported Tutee Samples, Treatment/Control Effects	72
Table 4.2	Descriptive Statistics for Continuous Variables, Treatment/Control Effects	73
Table 4.3	Descriptive Statistics for Categorical Variables, Treatment/Control Effects	74
Table 4.4	Method Descriptive Statistics by Year of Publication, Treatment/ Control Effects	75
Table 4.5	Description of Effects Used in the Treatment/Comparison Meta-Analysis	76
Table 4.6	Results from Research Question #1	91
Table 4.7	Results from the Research Question #2: Meta-Regression, Continuous Variables	92
Table 4.8	Results from the Research Question #2 Meta-Regression, Categorical Variables	92
Table 4.9	Coefficient Results from Multiple Meta-Regression Model #2	96

LIST OF FIGURES

Page

Figure 3.1 Search and Screening Process, 1960-2020 58

Figure 3.2 Comparing Classical and Multilevel Meta-Analysis 60

Figure 4.1 Hedge's g Effect Sizes for all Treatment/Control Outcomes 82

Figure 4.2 Outcomes for Treatment/Control Effects in Fluency..... 83

Figure 4.3 Outcomes for Treatment/Control Effects in Comprehension 84

Figure 4.4 Graphical Display of Study Heterogeneity 86

Figure 4.5 Funnel Plot of Observed Outcomes 88

CHAPTER 1

INTRODUCTION

Decades of research clearly link reading fluency achievement with reading comprehension outcomes (Fuchs et al., 2001; Kuhn et al., 2010; Kuhn & Stahl, 2003; National Reading Panel [NRP], 2000; Schwanenflugel & Benjamin, 2017; Teale et al., 2020; Therrien, 2004). The influence of fluency on comprehension is described by reading researchers using terms such as "bridge" (Pikulski & Chard, 2005, p. 511) "indicator" (Fuchs et al., 2001, p. 239), and "facilitat[or]" (Kuhn et al., 2010, p. 240). Fluent reading indicates efficient text processing that allows cognitive resources to attend to the construction and integration processes of comprehension (Kintsch, 2018; LaBerge & Samuels, 1974). Put briefly, smooth, accurate, and expressive reading helps create meaning from text (Kuhn et al., 2010).

The role of reading fluency in developing proficient readers cannot be underestimated. Students who read fluently tend to exhibit concomitant levels of comprehension achievement (Buck & Torgesen, 2018; Petscher & Kim, 2011; Smith et al., 2020). Although reading fluency does not guarantee comprehension, its absence will likely preclude the construction of meaning (Sabatini et al., 2019; Smith et al., 2020). Promoting proficient reading fluency, then, should be viewed as a major milestone in the development of all readers and an important crux of early reading instruction.

Although once referred to as the "neglected reading goal" (Allington, 1983, p. 556), reading fluency has received increased attention in the literature during the past two decades (NRP, 2000; Teale, 2020). This body of research confirms that oral fluency is a

milestone enroute to becoming a good reader, a notion the elementary education community has vigorously embraced (Hasbrouck & Tindal, 2017). The rise of curriculum-based fluency measures such as the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; University of Oregon., 2018-2019; Good et al., 2002) and Acadience (Good et al., 2011) has given teachers efficient methods to better screen and facilitate fluency development. Increased attention to reading fluency appears to correspond with nationwide fluency achievement as recent oral reading fluency norms by Hasbrouck and Tindal (2017) indicate rate increases for nearly all elementary grades.

Dysfluent Readers

Despite recent attention to fluency development as part of literacy instruction, many readers cannot read text fluently (Buck & Torgesen, 2018; Sabatini et al., 2019; Smith et al., 2020). Sabatini et al. estimated that 20% of fourth-grade students lack the requisite fluency skill needed to comprehend text; an alarming 600,000 fourth-grade students in the U.S. The cost of dysfluency is high; students who do not attain proficient oral fluency in elementary school are unlikely to become fluent in the secondary grades and further unlikely to reach proficient levels of reading comprehension (Paige et al., 2012). Such students will continue their schooling unprepared for frequent encounters with rigorous text, with consequences extending into adulthood (Mellard et al., 2012; Rasinski et al., 2017; Whithear, 2011). The plight of dysfluent readers should demand the utmost attention of literacy scholars and reading teachers.

Tutored Support in Reading Fluency

One technique used to promote oral reading fluency in weak readers is one-on-one tutoring (Dufrene et al., 2010; Hewison & Tizard, 1980; Rasinski & Stevenson, 2005). Several meta-analyses have indicated that tutoring delivered by minimally trained volunteers or stronger peers augments academic achievement across different content areas and tutee grade levels (Nickow et al., 2020; Ritter et al., 2009; Slavin et al., 2011). Other reviews have analyzed literacy-specific outcomes for one-on-one tutored support, and report effect sizes ranging from 0.29 (Cohen et al., 1982) to 0.41 (Elbaum et al., 2000; Leung, 2019). These meta-analyses indicate that tutoring has an overall positive effect on academic and literacy achievement. Beyond these initial conclusions, however, much less is known about the effects of specific tutoring regimens, or how specific tutoring regimens influence different types of reading outcomes such as oral reading fluency and comprehension.

Paired Oral Reading Methods

Several forms of tutoring involve pairing a weaker reader with a higher-level peer or adult reader to chorally read connected text together (Flood et al., 2005 A. Morgan et al., 2000; Tymms et al., 2011). These methods of Paired Oral Reading are prevalent in the literature (Kuhn & Stahl, 2003; Rasinski & Hoffman, 2003) using various terms such as Neurological Impress Method (NIM; Heckelman 1969), Paired Reading (Topping & Lindsay, 1992b), Dyad Reading (Eldredge & Butterfield, 1986), and Read Two Impress (Young et al., 2015). Evidence suggests that Paired Oral Reading techniques are effective for promoting a tutee's reading rate (Brown et al, 2018; Flood et al., 2005; Young et al.,

2015), oral reading accuracy (Topping & Lindsay, 1992a), oral reading prosody (Young, Durham et al., 2018), and silent reading comprehension (Downs et al., 2020; Flood et al., 2005). Table 1.1 outlines key details about the various methods of Paired Oral Reading.

Table 1.1

Description of Paired Oral Reading Methods

	Method			
Characteristic	Neurological impress method	Paired reading	Dyad reading	Read Two Impress
Original researcher	Rod Heckelman	Roger Morgan	J. Lloyd Eldredge	Chase Young
Key publications	Heckelman, 1967; Cook et al., 1981; Flood et al., 2006	Topping 1990; Topping & Lindsay 1992a	Eldredge & Quinn 1988; A. Morgan et al., 2000	Mohr et al., 2015; Young et al., 2015; Young, Pearce, et al., 2018
Primary tutor	Adult	Adult and Peer	Peer	Adult

Reports indicate that scholar and practitioner interest in Paired Oral Reading methods has waxed and waned throughout recent decades (Flood et al., 2005; Topping & Lindsay, 1992b). A recent surge of published research in the past decade suggests these practices are in vogue once again. Young et al. (2015, 2018a, 2018b, 2020) investigated stacking Paired Oral Reading with repeated reading (Samuels, 1979) to support readers who achieved below-benchmark proficiency and further work by Young and Rasinski (2017) described Paired Oral Reading as an intensive Tier 3 intervention using the Response to Intervention (RtI; Gersten et al., 2008) framework. Other recent studies investigated Paired Oral Reading in entire classrooms using higher achieving peers to tutor fellow third-grade students (Brown et al., 2018; Downs et al., 2020). Research

during the past decade also investigated Paired Oral Reading in diverse populations, such as foster children in the U.K. (Fry, 2014; Gately, 2014), English language learners in the U.S. (Klvacek et al., 2017, 2019; Vo, 2011) and India (Shah-Wundenberg et al., 2013), and students with Emotional Behavioral Disorders (Thornton, 2012). This recent scholarly interest in Paired Oral Reading appears to extend into practice; one state recently included Paired Oral Reading as a major component in a curriculum that utilizes community volunteer tutors to assist weaker readers in elementary grades (Utah State Board of Education [USBE], 20219.

Clearly, promoting reading achievement through Paired Oral Reading is experiencing renewed interest. Less clear however, are the specific factors driving the current surge. Increased emphasis in areas such as oral reading fluency achievement (NRP, 2000), or reading intervention using the Response to Intervention framework (RtI; Gersten et al., 2008) could be influencing the current comeback. Other factors include the current interest in supporting students in reading challenging text, influenced by the widespread adoption of the Common Core State Standards (CCSS; National Governors Association 2010), or the predictive value of fluency on end-of-level reading assessments (Smith et al., 2020). Regardless of specific influencers, interest in Paired Oral Reading methods is perhaps at an all-time high, with implications for research and practice.

Problem Statement

Despite decades of research and purported popularity in the classroom, the effects of Paired Oral Reading practices on student literacy outcomes, specifically fluency and comprehension, are not yet well understood. The extant literature does not indicate

optimal intervention lengths or dosages, nor does it suggest which variations of Paired Oral Reading might be most effective, or which tutee ages or levels of reading achievement could most benefit. Indeed, the extant literature on Paired Oral Reading methods appears fractured among four separate methods; each technique remains isolated from the other techniques. For example, Downs et al. (2020) recently implemented Dyad Reading with third-grade students for half a school year. These authors referenced and explicated findings within the scope of previous research in Dyad Reading (c.f., Brown et al., 2018; Eldredge & Quinn, 1988; A. Morgan et al., 2000), but made no attempt to synthesize results of other methods of Paired Oral Reading (c.f., Flood et al., 2005; Tymms et al., 2011; Young et al., 2015, 2018a, 2018b). Such siloed reporting–typical among the various methods of Paired Oral Reading–could constrict the further development of each technique.

Further, despite six decades of existence, very minimal work attempts to synthesize literacy achievement outcomes via systematic review among all methods of Paired Oral Reading (cf., Kuhn & Stahl, 2003; Rasinski & Hoffman, 2003; Topping & Lindsay, 1992a). No known systematic reviews or meta-analyses of Paired Oral Reading outcomes exist. The lack of synthesis renders the relative effect of the practice and the influence of key moderating variables unknown, confounding any coherent future research agenda of Paired Oral Reading. Ultimately, a lack of clarity within the Paired Oral Reading literature may influence practical implementations of Paired Oral Reading in classroom settings.

Purpose of the Meta-Analytic Study

Given the current lack of consolidation of related research this study sought to determine average effects of Paired Oral Reading methods. Specifically, the study meta-analyzed studies that measured fluency and comprehension outcomes for a peer or adult tutor reading aloud synchronously in connected text with a lower-level reader. Meta-analytic methods synthesize outcomes from a range of related studies to calculate an average estimated effect (Bus et al., 2011). Two major advantages of meta-analysis include investigating specific factors that enhance or inhibit outcomes reported in the literature (Card, 2017; Tipton et al., 2019b), and controlling for these variables in the effect size calculation using multiple meta-regression.

Meta-analytic techniques have evolved rapidly in the past decade to include a new generation of robust methods that can better estimate outcomes and control for moderating variables (Fernández-Castilla et al., 2020). Among these techniques, multilevel regression meta-analysis affords particular strengths for literacy researchers. Foremost, a multilevel analysis allows for multiple effect sizes from a single study to be analyzed. This affordance is facilitated by including an intermediary level of regression that directly models the covariance between multiple effect sizes within a single study. (Hox et al., 2018). A multilevel, multivariate meta-analytic design is beneficial for literacy research because this design more accurately reflects the design of literacy studies. Quantitative literacy research tends to report multiple outcomes, whether through the use of multiple treatment groups, multiple measures, or both. A multilevel meta-analytic design is better equipped to incorporate these designs in the analysis (Fernández-

Castilla et al., 2020).

Within the current context, the multilevel meta-analysis of Paired Oral Reading calculated the average outcome of these techniques via effect size and determined specific factors that influenced outcomes. Critically, this meta-analysis sought to parse any differential effect of Paired Oral Reading techniques on fluency and comprehension outcomes. Results of the analyses provided a more nuanced understanding of Paired Oral Reading methods, informing future research and classroom practice. Similarly, outcomes from this meta-analysis highlight the viability of future multilevel techniques in literacy research.

Research Questions

Using multilevel meta-analytic methods, this study addresses the following questions.

1. What is the effect of Paired Oral Reading on reading outcomes for tutees in grades K-6?
2. How do related factors (i.e., tutor type, dosage, year, publication type, and method used) moderate the reported outcomes?
3. What are the effects of Paired Oral Reading for tutees while accounting for significant moderator variables?

Limitations

This meta-analytic study is limited by various factors. Perhaps most prominent, the quality of the literature corpus used in the analysis likely limited the methodological design of the study and the reported outcomes. Meta-analyses provide an estimated effect of the true outcome (Borenstein et al., 2010), and thus less rigorous studies likely

influenced the accuracy of reported outcomes. Further, this study was limited by the volume of literature meeting the inclusion criteria. Like any statistical analysis, meta-analysis relies on adequate, representative sampling to calculate reliable outcomes. In the case of meta-analysis, the corpus of studies themselves serve as the research sample. Nonsignificant results in a meta-analysis can be the result of inadequate power, "rather than the absence of real effect in the population" (Hox et al., 2018, p. 230). Moreover, the reported outcomes may be limited by efficacy and accuracy of the multilevel meta-analytic methods used (Hox et al., 2018).

To mitigate these limitations to the greatest extent possible, concerted effort was made to analyze the extant data with innovative multilevel analysis techniques. Use of multilevel meta-analysis has expanded rapidly in the last five years (Fernández-Castilla et al., 2020), and is favored for hierarchal data sets because it produces more accurate estimates than single level analyses (Hox et al., 2018).

Delimitations

As is typical in synthetic research, some constraints have been applied to increase the integrity of the study. After careful review of the extant literature, certain types of related practices were deemed to lie outside the scope of this study. Specifically, the meta-analysis only investigated one-on-one tutoring schemes that pair a lower-level reader with a higher-level reader in the elementary grades. Other similar tutored reading schemes were excluded, namely synchronous audio assistance (Esteves & Whitten, 2011; Hollingsworth, 1970; Mikkelsen, 1981), triad choral reading (Eldredge, 1990), preschool-age shared book reading (Dixon-Krauss et al., 2010; Seitz & Bartholomew, 2014), Paired

Oral Reading as a minor component of a larger, structured intervention (Rasinski, 1994; Wilfong, 2008), peer-assisted learning strategies (PALS; Fuchs et al, 2000; Mathes et al., 2003) and other forms of tutored, non-synchronous reading assistance (Gallagher, 2008; Lancy & Nattiv, 1992; Medcalf, 1989; Shany & Biemiller, 1995). Further, Paired Oral Reading occurring in a secondary setting (King, 1986; Lloyd et al., 2015) or at home (Leach & Siddall, 1990; Law & Kratochwill, 1993; A. Miller et al., 1986; B. V. Miller & Kratochwill, 1996) was excluded in this analysis. Further, this analysis investigated adult, cross-age, and peer-tutored tutee outcomes. Due to this structure, investigation of the outcomes of peer-tutors was excluded. Exclusion of these criteria afforded a more precise calculation of the effect of Paired Oral Reading for elementary students in a school environment.

Conclusion

Oral reading fluency is considered a critical component of successful reading development, yet evidence suggests many elementary readers remain dysfluent. Recently, researchers and practitioners alike have resurrected methods of Paired Oral Reading to promote the reading outcomes of young readers. However, research efforts targeting these methods are isolated, obfuscating their effect on student literacy outcomes. The following systematic review of Paired Oral Reading techniques and subsequent multilevel meta-analysis provides valuable insights about the reading achievement outcomes for tutees engaged in these practices.

Definition of Terms

Reading Tutoring: Support delivered by one tutor to one tutee, designed to improve the reading achievement of the tutee in one or more specific areas (Shanahan, 1998).

Fluency: The combination of "accuracy, automaticity, and oral reading prosody, which, taken together, facilitate the reader's construction of meaning" (Kuhn et al., 2010, p. 240).

Comprehension: "The process of simultaneously extracting and constructing meaning through interaction and involvement with written language" (Snow et al., 2002, p. xiii).

Paired Oral Reading: A series of related practices where a higher-level reader (adult or peer) synchronously reads connected text with a lower-level reader. The higher level and lower level in the Paired Oral Reading literature are commonly referred as tutor and tutee (Topping & Lindsay, 1992b), or lead and assisted reader (Eldredge & Quinn, 1988).

Neurological Impress Method (NIM; Heckelman, 1969): A Paired Oral Reading technique where an adult tutor and student tutee read connected text aloud, synchronously.

Paired Reading (R. Morgan & Lyon, 1979): A Paired Oral Reading technique where a tutor (peer or adult) alternates between synchronous read aloud and tutee independent read aloud (R. Morgan & Lyon, 1979).

Dyad Reading (Eldredge & Quinn, 1988): A variation of NIM where a peer tutor

synchronously reads aloud with a student tutee.

Read Two Impress (Young et al., 2015): A stacked variation of NIM that combines repeated reading (Samuels, 1987) with synchronous oral reading between an adult tutor and a student tutee.

Meta-analysis: A form of systematic review involving the "quantitative analysis and synthesis of a set of related empirical studies in a well-defined domain" (Bus et al., 2011, p. 270).

Classical Meta-analysis: Methods of random-effect and fixed-effect meta-analysis that utilize a two-level hierarchical structure to account for outcomes nested within studies. (Cheung, 2014; Hox et al., 2018).

Multilevel Meta-analysis: Meta-analyses that use three or more hierarchical levels to model covariances within study data (Fernández-Castilla et al., 2020; Hox et al., 2018). The standard three-level multilevel meta-analysis models outcome-level effect sizes, nested within covariance between effect sizes, nested within covariance between studies (Fernández-Castilla et al., 2020; Van den Noortgate et al., 2015). Multilevel meta-analysis allows the statistical model to reflect the extant structure of the data more accurately, allowing for a more precise analysis. Further, adding the intermediate second level allows for a multivariate analysis to account for the multiple outcomes from a single study.

CHAPTER 2

LITERATURE REVIEW

Variations of Paired Oral Reading are currently experiencing renewed interest in academic research and practitioner settings (Downs et al., 2020; USBE, 2020; Young et al., 2015). These new iterations reflect the hope that teachers have that daily oral reading between partners is well worth the time allotted and the belief that there are positive reading and social outcomes. Although tutoring is generally viewed as an effective method to promote reading achievement in weaker readers (Cohen et al., 1982; Nickow et al., 2020; Shanahan, 1998), the specific effects of Paired Oral Reading practices on participants remain unknown. A thorough literature review and subsequent meta-analysis may assist researchers and practitioners alike to better understand this popular approach to supporting oral reading and enhance its application in clinical and classroom environments, which are constrained by time and resources.

Search Process

This literature review of Paired Oral Reading techniques included an extensive search process to (a) identify lines of research that have targeted Paired Oral Reading in support of reading achievement and (b) identify studies that included fluency and comprehension among the outcomes for students who participated. These objectives were achieved through a database search conducted in Academic Search Ultimate, APA PsychInfo, ERIC, OpenDissertations, and Professional Development Collection using various terms to identify methods of Paired Oral Reading between 1980-2020. The search

included the following terms: *dyad read*, neurologic* impress method, "pair* read*," partner read*, "peer-assisted reading," "prime-o-tec," "reading-while listening," "listening while reading," "shadow read*," "shar* read*," "assisted reading," "duet read*," "duolog read*," paired assisted read*, paired partner read*, impress read*, read two impress, and "buddy read*."* The resultant review of literature subsequently informed the inclusion/exclusion criteria that determined the specific studies to include in the meta-analysis. In other words, a broad range of relevant scholarship informed the researcher's understanding of the topic; however, a more limited set of studies afforded metrics eligible for review.

Categorizing Paired Oral Reading Techniques

The extant literature demonstrates the prevalence of various Paired Oral Reading techniques throughout the past five decades (Flood et al., 2005; Kuhn & Stahl, 2003; Rasinski & Hoffman, 2003; Topping, 2017; Topping & Lindsay, 1992a). These studies report various techniques that involve a higher-level reader and lower-level reader reading connected text aloud in unison. Despite this unifying trait, the Paired Oral Reading literature presents two major differences in application. First, some research includes the tutoring pair alternating between synchronous reading and independent tutee reading (Topping & Lindsay, 1992a; Young et al., 2015), whereas other research includes synchronous reading through the duration of the session (Downs et al., 2020, McAllister, 1989). The second major difference involves the use of adult tutors (Flood et al., 2005) versus peer tutors (Brown et al., 2018).

History and Evolution of Paired Oral Reading Methods

First Wave: Origin of Paired Oral Reading Methods (1966-1979)

Separate methods of Paired Oral Reading emerged in the U.S. and the U.K. within a decade of each other (Heckelman, 1966; R. Morgan, 1976). Despite originating in two different countries with the originators seemingly unaware of each other, these methods share striking similarities. First, these methods consist of extended synchronous oral reading of authentic text with an adult tutor. Second, these methods sought to remediate the choppy, halting, "phonics-bound condition" (Heckelman, 1969, p. 281) of dysfluent readers through direct tutelage (J. M. Barrett, 1987; R. Morgan, 1976; Topping, 1990). Third, clinical psychologists—not academic researchers—pioneered both methods. Fourth, both methods focused on the tutor and tutee reading as much text as possible in each session, with no attention toward decoding or other isolated skill instruction. Fifth, both methods quickly evolved to include lay volunteers trained to be tutors.

U.S. Origins–Neurological Impress Method

Heckelman (1966, 1969) believed that poor readers' brains possessed faulty reading networks that impeded correct patterns of reading (Embrey, 1968). After learning that stuttering could be remediated if individuals listened to their own speech while speaking, Heckelman (1986) hypothesized that reading difficulty could be remediated in a similar manner. The result was a method of Paired Oral Reading where an adult tutor reads aloud continuously with a weak reader during an entire intervention session. Purportedly, the effective reading patterns of the tutor could be "deeply impressed"

(Embrey, 1968, p. 33) upon the tutee, assisting the development of "fluidness" (Heckelman, 1969, p. 238). Heckelman (1966, 1969) initially investigated this approach with several poor readers in a clinical setting and reported impressive results. Coined Neurological Impress Method, the pragmatic nature and early success of the method led to studies conducted at elementary schools with adult tutors (Embry, 1968; Lorenz & Vockell, 1979).

Heckelman (1969) apparently did not craft his method of Paired Oral Reading using a clear theoretical framework. Some evidence suggests he worked from a loosely defined cognitive paradigm; he posited Paired Oral Reading provided a correct model of reading that over time becomes "deeply impressed" (p. 282) on the neurological systems of the tutee. Although not explicitly noted as a cognitive approach by Heckelman, this initial framing of Paired Oral Reading may have served as a precursor to the cognitive perspectives that would later dominate the Paired Oral Reading literature.

U.K. Origins–Paired Reading

Seemingly unaware of Heckelman's recently developed method, R. Morgan (1976) crafted another method of Paired Oral Reading for clinical use in England. Morgan also viewed reading difficulty as a performance deficit best remediated by extended practice with connected text, rather than practicing isolated skills. However, inspired by the principles of operant conditioning (Skinner, 1937), Morgan developed a method of Paired Oral Reading that relied on specific tutor reinforcements. The approach represented a behaviorist model because proficient tutee reading was reinforced via behavioral prompts by the tutor. The method, called Paired Reading purportedly

reinforced reader development by providing continual accurate prompts (synchronous reading), immediate error correction, and periodic praise for correct responses (R. Morgan & Lyon, 1979). The approach represented a behaviorist model of Paired Oral Reading; proficient tutee reading was reinforced via behavioral prompts by the tutor. Notably, this method differed from Heckelman's (1969) Neurological Impress Method because tutor and tutee alternated between reading synchronously and tutee independent reading, based on tutee non-verbal signals and word-reading errors.

Morgan intended the new method to be simple enough for lay tutors to implement. Initial studies successfully trained paraprofessionals and parents as tutors, attracting national interest in the U.K. (R. Morgan 1976; R. Morgan & Lyon 1979). Likely bolstered by widespread reform efforts to improve academic outcomes (Hewison & Tizard, 1980; A. Miller, 1985; Spalding et al., 1984), R. Morgan's (1976) method of Paired Oral Reading later became a popular tutoring technique in the U.K. during the 1980s. Parents were most commonly recruited as tutors and encouraged to use Paired Oral Reading with their readers at home (J. M. Barrett, 1986; Bush, 1983; Lindsay et al., 1985; A. Miller et al., 1986; Spalding et al., 1984).

First Wave Conclusion

The first wave of Paired Oral Reading research included the conception of two techniques that utilized adult tutors to read connected text aloud synchronously with a tutee. Although seemingly founded on different theoretical premises, these techniques shared a pragmatic structure to support readers facilitating their subsequent popularity adaptation from clinical to school settings.

Second Wave: Theoretical Evolutions and Tutor Adaptations (1980-1991)

Evolution from the Behaviorist Stance

While the Neurological Impress Method remained mostly atheoretical, Paired Reading distanced itself from the behaviorist perspective during the 1980s. One implication of a behaviorist view of Paired Oral Reading was that variance among tutor feedback should correlate with tutee outcomes; tutees with more frequent and accurate reinforcers or corrections during reading should achieve better reading outcomes (R. Morgan & Lyon, 1979). This hypothesis was not evidenced in the research; several studies found that tutors who only loosely followed the behaviorist protocol still produced effective results (Limbrick et al., 1985; Winter, 1988). Winter asserted that "the data fail to reveal any relationship between the behaviour of tutors and tutees during sessions and the reading gains made by tutees at project end" (p. 147). These findings, combined with surging interest in top-down cognitive theories, led to the decline of behaviorist explanations of Paired Reading.

The 1980s saw the rise, and later decline, of whole language and self-esteem theoretical frameworks of Paired Reading (J. M. Barrett, 1987; Bushell et al., 1982; Diaper 1989; A. Miller et al 1986; Topping, 1990). The whole language position favored extended encounters in authentic texts, such as the tutoring provided by Paired Oral Reading. Purportedly tutor support in Paired Oral Reading developed a tutee's semantic and syntactic processing, promoting oral language and overall reading achievement. This framework was frequently used to explain reader progress resultant from Paired Oral Reading, however the whole language paradigm itself was rarely examined in these

studies (cf., Joscelyne, 1991).

Similarly, self-esteem frameworks were frequently invoked throughout the 1980s, but rarely examined. This position contended that low self-esteem was the predominant barrier to reading progress, perhaps due to a history of reading difficulty (Bushell et al., 1982). Allegedly, a committed tutor could promote reading development by promoting a desire to persist through text and cultivate a sense of success in the reader (A. Miller 1985; Winter, 1988). Hypothetically, the reader's self-esteem would subsequently grow through successful tutoring experiences, that in turn would influence reader achievement. This view held that greater emotional support from a tutor could increase reading outcomes of the tutee even further. Although popular during 1980s, the whole language and self-esteem frameworks that predominantly framed Paired Reading techniques were never well established and not empirically measured (Joscelyne, 1991). Joscelyne robustly analyzed and critiqued these explanatory perspectives, marking their de facto end.

Rise of Text Processing Theoretical Frameworks

During the late 1980s, bottom-up frameworks became the dominant premise of Paired Oral Reading, beginning with the Neurological Impress Method. Initially, Henk (1983) and Kann (1983) noted potential text processing advantages afforded to tutees through Paired Oral Reading. Heckelman (1986) later clarified his original (1969) position by stating that the Neurological Impress Method "speed[s] up of the processing of information [which] reinforces…the automated information acquisition process" (1986, p. 415). These initial speculations appear to indirectly tie Paired Oral Reading

with LaBerge and Samuels' (1974) theory of automaticity.

Eldredge (1988a, 1988b, 1988c, Eldredge & Quinn, 1988) was the first to link Paired Oral Reading directly and consistently to LaBerge and Samuels' (1974) theory of automaticity, which asserts that extended exposure to reading practice helps readers develop the ability to recognize words automatically. Word reading automaticity purportedly consumes relatively fewer cognitive resources, allowing attention to be directed toward other aspects of reading such as oral expression and comprehension (Samuels, 1987). Within the context of Paired Oral Reading, Eldredge and Quinn noted that Paired Oral Reading likely "help[s] poor readers focus on the important aspects of text, free[s] them from the decoding burden, and speed[s] up the decoding process so they can give necessary attention to text message" (p. 33). Echoing Eldredge's sentiments, U.K. researchers (Topping 1990; Topping & Lindsay, 1992b) lent further support to the automaticity position, proposing that Paired Oral Reading "frees the struggling reader from a preoccupation with laborious decoding and enables other reading strategies" (Topping, 1990, p. 15).

The 1980s represents a dynamic period for the theoretical understanding of Paired Oral Reading methods. The decade opened with behaviorist perspectives dominating the research. This perspective was later challenged by top-down and self-esteem theories of reader achievement, which were in turn replaced by the bottom-up cognitive theory of developing automatic text processing. Although theoretical squabbles occasionally appear after this time period (c.f. Topping, 1997; Winter, 1996, 1997), the Paired Oral Reading literature has consistently adopted a theory of automaticity after the late 1980s

(Almaguer, 2005; Brown et al., 2018; Downs et al., 2020; Flood et al., 2005; Topping, 2017; Young et al., 2015). Theoretical innovations were not the only major contribution to the literature during this decade; tutor adaptions would allow Paired Oral Reading to reach a much broader audience.

Introduction of Peer Tutors–Peer Tutor Paired Reading and Dyad Reading

In addition to the significant theoretical evolutions of the 1980s, the last half of that decade included research utilizing same- and cross-age peers to serve as Paired Oral Reading tutors. (Eldredge & Butterfield, 1986; Eldredge & Quinn 1988; Limbrick et al., 1985; Low & Davies, 1988; Low et al., 1987; McAllister, 1989). Initial peer-tutor studies conducted by U.K. Paired Reading researchers demonstrated that Paired Oral Reading methods are simple enough to be implemented by a peer (Limbrick et al., 1985; Low & Davies, 1988). Concurrently, researchers in the U.S. adapted the Neurological Impress Method to include peer tutor variations (Eldredge & Butterfield 1986; McCallister, 1989)

Eldredge and colleagues (Eldredge & Butterfield 1986; Eldredge & Quinn 1988) adopted a unique perspective for their iteration of Paired Oral Reading. These researchers perceived Paired Oral Reading as a method to provide weak readers with support during basal and classroom instruction. Their technique, termed Dyad Reading, integrated Paired Oral Reading with regular classroom instruction, facilitating a poor reader's access to basal content and subject-area instruction while promoting reader development. Weak readers were provided Paired Oral Reading assistance throughout the day as needed with peer tutors regularly rotating in and out of service. This adaptation conceivably varied the

dose of Paired Oral Reading in accordance with student need. Further, students were released from participation in the method once they achieved the expected proficiencies.

The stance adopted by these researchers represents a significant, albeit temporary, shift in the literature. Previous research implemented Paired Oral Reading within the school day for a predetermined amount of time, separate from regular classroom instruction (Limbrick et al., 1985; Low & Davies, 1988; McAllister, 1989). Eldredge and colleagues however sought to integrate Paired Oral Reading across reading instruction and activities throughout the school day. This approach yielded positive results; two studies (Eldredge & Butterfield 1986; Eldredge & Quinn 1988) reported significant outcomes in vocabulary and comprehension using the Gates-MacGinitie Reading Test for participant tutees. Unfortunately, these innovative and pragmatic applications were never further investigated.

Second Wave Conclusion

The second wave of Paired Oral Reading is marked by heightened popularity, theoretical evolutions, and peer adaptations adaptions. The period throughout the 1980s saw Paired Reading researchers abandon the behaviorist stance, and subsequently evolve through whole language and affective paradigms before adopting a text processing framework. Neurological Impress Method avoided much of this turbulence and tangentially adopted a text processing framework during this period. The second wave of Paired Oral Reading research also exhibited the utilization of peer tutors in Paired Oral Reading research. The utilization of text processing frameworks to explain Paired Oral Reading outcomes and the use of peer tutors would later become a staple of Paired Oral

Reading research.

Waning Popularity (1993-2000)

Compared to the bustling activity associated with Paired Oral Reading research throughout the 1980s, the 1990s represent a relative lull in the Paired Oral Reading literature. Several studies during this era critiqued the fidelity of parent-directed tutoring schemes (DeAngelo et al., 1997; Law & Kratochwill, 1993; B. V. Miller & Kratochwill, 1996; A. L. Miller & Narrett, 1995). These studies subsequently focused on developing more effective training and accountability of parent tutors. Other reports involve minor critiques and counter-critiques of previously conducted research in the U.K. (Topping, 1997; Winter, 1996, 1997). The reasons for the sudden decline are not reported in the literature, however the first decade of the 21st century would see a gradual increase in Paired Oral Reading interest.

Third Wave: Contemporary Paired Oral Reading Research (2000-Present)

Gradual Resurgence

Paired Oral Reading research in the 21st century represents re-emergence of these techniques and subsequent evolution into more sophisticated and nuanced investigations. A. Morgan et al. (2000) initiated this resurgence by investigating the role of text difficulty on tutee outcomes using Dyad Reading. This study would later influence a pair of replication studies (Brown et al., 2018; Downs et al., 2020) designed to better understand text difficulty as a variable. Later, Tymms et al. (2011) reported on Paired Reading using a randomized control trial design. Soon thereafter, a trio of studies

continued to evolve Paired Oral Reading research into new areas of study. Almaguer (2005) investigated Dyad Reading with English language learners and Flood et al. (2005) incorporated a comprehension component into the Neurological Impress Method using adult volunteer tutors. That same year, Canadian researchers Sokal et al. (2005) investigated Paired Reading with primarily Aboriginal boys in inner-city Winnipeg. The research conducted during this decade appears to represent a renewed interest in these techniques and implementation with different populations and for additional purposes.

Increased Popularity

Additions to the Paired Oral Reading literature appear to have accelerated during the past decade. Fry (2014) and Gately (2014) utilized the Paired Reading technique to support foster learners in the U.K. Young et al. (2015, 2016, 2020; Young, Durham, et al., 2018; Young, Pearce, et al., 2018) provided a stacked intervention, termed Read Two Impress, integrating the Neurological Impress Method with Samuels' (1987) method of repeated reading. In a comparative study, the combined approach of Read Two Impress was found to outperform the Neurological Impress Method on measures of reading fluency and reading comprehension with elementary age students (Young, Pearce, et al., 2018). Several recent studies also investigated the influence of Paired Oral Reading on measures of tutee prosody using peer (Brown et al., 2018) and adult tutors (Young, Durham, et al., 2018). Acknowledging the pragmatic nature of these practices, Young and Rasinski (2017) specifically recommended utilizing Paired Oral Reading as an intensive intervention within the Response to Intervention (RtI; Gersten et al., 2008) framework.

A consistent focus on text complexity has also emerged in the past five years (Brown et al., 2018; Downs et al., 2020; Young, Pearce, et al., 2018; Young, Durham, et al., 2018). Two of these studies (Brown et al., 2018; Downs et al., 2020) invoked findings from A. Morgan et al. (2000) to investigate how text complexity influenced fluency and comprehension outcomes with third-grade peer tutors using Dyad Reading. Curiously, these studies noted 'surges' (Brown et al., 2018, p. 550) in student proficiency over the duration of the study. The surges were characterized by periods of accelerated achievement among peer tutees, which later attenuated. Both studies demonstrated tutees reading texts above their instructional level with successful outcomes.

Concern for reader attitudes is also a focal point of this time period, with studies investigating the influence of Paired Oral Reading on general reading attitudes, (Brown et al., 2018; Downs et al., 2020; Young, Pearce, et al., 2018) and attitudes toward peer-assisted Paired Oral Reading itself (Downs et al., 2020). Exact reasons for this resurgence of attitudes in the literature are not well delineated, however Downs et al. cite concerns about one-size-fits-all implementations and widespread practice without attention to dosage. Brown et al. (2018) noted a positive response from peer tutors and tutees toward Dyad Reading at the end of the intervention, however attitudes were not formally measured. Downs et al. (2020) subsequently hypothesized that the social collaboration inherent in partnered reading may positively influence reading attitudes, and affective measures could be included to monitor reader identity. When formally measured however, reading attitudes did not appear to significantly increase as a result of whole-class Dyad Reading (Downs et al., 2020). These null findings for reader attitudes were

similar to an earlier finding from Young, Pearce, et al. (2018) study using the Read Two Impress method.

The current zeitgeist in Paired Oral Reading also represents a marked shift in study design and rigor. Recent study designs in Paired Oral Reading utilize multiple treatment groups (Brown et al., 2018; Young, Pearce, et al., 2018), random assignment (Brown et al., 2018; Young, Pearce, et al., 2018) increased sample sizes (Brown et al., 2018, Downs et al., 2020), and multilevel analysis (Brown et al., 2018, Downs et al., 2020, Lloyd et al., 2015; Tymms et al., 2011). These evolutions are a welcome addition to the predominantly small sample sizes and simple analyses utilized in earlier research (c.f., Cook et al., 1980; Topping & Whitely, 1993; Winter, 1988).

Common Features of Paired Oral Reading

Tutor Variations

Understanding the influence of different tutor types on reading outcomes should be an important goal for Paired Oral Reading researchers. Tutors are requisite for any form of Paired Oral Reading; however, tutor types vary across studies and methods. Same-age peers, cross-age peers, and adults have all performed the tutoring role as noted in the extant scholarship. Adult tutors are noted in the literature to be typically volunteers (Young, Durham, et al., 2018), parents (Law & Kratochwill, 1993), preservice teachers (Flood et al., 2005), or paraprofessionals (Flood et al., 2005; R. Morgan, 1976). Studies involving same-age peers generally include entire classes divided into tutor and tutee roles (Brown et al., 2018; Downs et al., 2020; Tymms et al., 2011) or specific students screened as below grade level (Eldredge & Quinn, 1988; Joscelyne, 1991; A. Morgan et

al., 2000; Young et al., 2015). Cross-age tutoring is less common in the extant literature, however two grade levels between tutor and tutee appears to be a preferred design when used. (Limbrick et al., 1985; Low et al., 1987; Tymms et al., 2011).

Grade Level of Tutees

Clearly, any successful implementation of Paired Oral Reading should involve a nuanced understanding of differences among readers at various grade levels. The extant literature however does not indicate optimal grade levels for Paired Oral Reading tutoring. Paired Oral Reading has been studied predominantly in elementary-age students (Cadieux & Bourdrealt, 2005; Downs et al., 2020; Eldredge & Quinn, 1988; Law & Kratochwill, 1993, Sokal et al., 2005; Topping & Bryce, 2004; Winter, 1988; Young et al., 2015). Research investigating Paired Oral Reading has occurred in every elementary grade, with second- and third-grade students the most common participants. A smaller set of studies investigated Paired Oral Reading in middle- and high-school settings (Bedsworth, 1991; Cawood & Lee, 1985; King, 1986; Robson et al., 1984; Topping, 1990).

Location of Intervention

This literature review identified research in Paired Oral Reading being conducted within school and home environments. Paired Oral Reading within a school environment is a consistent feature of the literature across the past four decades (Downs et al., 2020; Eldredge & Quinn, 1988; Flood et al. 2005; A. Morgan et al., 2000; McAllister, 1989; Limbrick et al., 1985; Young et al., 2015). These studies typically trained adult

volunteers or peers to serve as tutors during the school day. The extant literature also contains research conducted within home environments, typically training parents to use R. Morgan's (1976) Paired Oral Reading technique with their children (J. M. Barrett, 1987; Bush, 1983; Leach & Siddall, 1990; Lindsay et al., 1985; A. Miller et al., 1986; R. Morgan & Gavin, 1988; Spalding et al., 1984). Home studies of Paired Oral Reading occurred primarily in the U.K. during the 1980s. Many of these home studies were published as conference reports (Cawood & Lee, 1985; Lees, 1985; O'Hara, 1985; Ripon et al., 1986; Simpson, 1985; Welsh & Roffe, 1985; Vaughey & MacDonald, 1986) with questionable methodological rigor, however some were published in peer-reviewed journals (J. M. Barrett, 1987; Limbrick et al., 1985; Lindsay et al., 1985; A. Miller et al., 1986). The current fourth wave of Paired Oral Reading research implements these schemes nearly exclusively within school settings (Brown et al., 2018; Downs et al., 2020; Young, Durham, et al., 2018; Young, Pearce, et al., 2018).

Intervention Duration and Frequency

An important variable in any effort to remediate a reading deficiency is the amount of time devoted to the intervening practice and necessary to achieve the desired outcomes. Determining the ideal amount of time spent in paired reading in individual sessions and across weeks of time are important goals for researchers. The duration of Paired Oral Reading interventions reported in the literature varies widely. Earlier studies describe interventions lasting as short as 4-6 weeks (Law & Kratochwill, 1993; Low et al., 1987; A. Miller et al., 1986), as long as six months (Strong & Traynalis-Yurek, 1983), and even an entire school year (Eldredge & Quinn, 1988). Studies published since

2000 have commonly reported durations of 4-7 weeks (Fiala & Sheridan, 2003; Gately, 2014; Young et al., 2015; Young, Pearce, et al., 2018) to approximately a half of a school year (Brown et al., 2018, Downs et al., 2020; A. Morgan et al., 2000, Tymms et al., 2011). Although session duration also varies, 15 minutes per session has been a common dosage across studies (Downs et al., 2020; Flood et al., 2005; Young et al., 2015).

In home and school environments, Paired Oral Reading interventions have been most commonly implemented daily (Brown et al., 2018; Leach & Siddal, 1990; Lindsay et al, 1985; A. Miller et al., 1986). Other frequency intervals include weekly (Ellis, 1995), twice weekly (Lees, 1985; Sokal et al., 2005), and thrice weekly (Young, Pearce, et al., 2018). A limited number of studies have investigated variations in session frequency for Paired Oral Reading. For example, Tymms et al. (2011) measured differences between weekly and thrice weekly Paired Oral Reading at school while B. V. Miller and Kratochwill (1996) measured a standard dose of 400 minutes for students who completed the dosage at various intervals across 5 months.

Text Selection and Text Difficulty

Paired Oral Reading involves tutor and tutee chorally reading connected text. The text thus plays an important role in these interventions. In the Paired Oral Reading literature, tutees most commonly self-select the texts used for oral reading practice. Interestingly, Heckelman (1969) and R. Morgan (1976; R. Morgan & Gavin, 1988)–the originators of Paired Oral Reading–advocated for advanced texts to be a major component of their methods. However, subsequent studies tended to neglect difficult texts in favor of allowing students complete autonomy in text selection (Cook et al.,

1980; McAllister, 1989; Winter, 1988). This small, but critical, adjustment appears to be supported by the rationale that full autonomy in text selection would provide tutees with a more engaging experience (A. Miller, 1985; A. Miller et al., 1986; A. Miller, 1987; Topping & Lindsay, 1992b), and was perhaps tied to theoretical shifts toward whole language and self-esteem frameworks (A. Miller et al., 1985). One study (Limbrick et al., 1985) noted that when tutees were allowed complete autonomy, they frequently selected books that were more difficult than what they could read independently, likely due to receiving tutor support. Joscelyne (1991) formally measured this notion and found that students selected text within their independence level regardless of tutor influence.

The last two decades represent a shift in the literature for the selection of texts. Since 2000, studies were more likely to include tutees reading texts above their instructional level as an explicit component of the intervention (Downs et al., 2020; Flood et al., 2005; Tymms et al., 2011; Young, Durham, et al., 2018). A pair of studies (Brown et al., 2018; A. Morgan et al., 2000) specifically investigated the role of text difficulty in Paired Oral Reading schemes. Both studies reported that students (second and third graders) reading two grade levels above the tutee's instructional level experienced the most progress on measured reading outcomes. Studies that utilize advanced texts have included an aspect of student self-selection (e.g., students choosing a book from a leveled bin; Brown et al., 2018; Downs et al., 2020); however, such choices are within the parameters of a predetermined difficulty level. Despite apparent neglect in the literature for decades, it appears that tutees reading texts above their instructional level has more recently become a standard of the Paired Oral Reading practices

Notably, a small corpus of studies minimized or did not allow for student self-selection of text. These studies involved students enacting Paired Oral Reading with a classroom reading basal or other teacher-selected reading materials (Eldredge & Butterfield, 1986; Eldredge & Quinn, 1988; Fiala & Sheridan, 2003; Low & Davies, 1988). However, participants in these studies were identified as below-grade level, and although not stipulated in the reports, were thus possibly reading texts above their instructional level.

Reported Academic Outcomes of Paired Oral Reading

Methods of Paired Oral Reading have been studied among young readers for decades. Critically however, these methods have been studied in isolation from one another. There is no known literature that attempts to synthesize the academic outcomes of Paired Oral Reading. The extant corpus consists of a literature review and two large-scale studies of a single method of Paired Oral Reading, and approximately 50 quasi-experimental, single-group designs, and reports of school-implemented projects conducted primarily in the U.S. and the U.K. These studies typically measure comprehension and fluency outcomes before and after a Paired Oral Reading intervention. Although research typically reports positive outcomes for readers involved in Paired Oral Reading, academic outcomes in the extant literature vary widely. The following synthesis presents the reported outcomes of Paired Oral Reading and highlights some of the complexities within the existing literature.

Literature Reviews and Large-Scale Studies

R. Morgan's (1976) Paired Reading technique has been communicated via several influential publications, including a report on 155 school implemented projects (Topping, 1990), a literature review covering the years 1976-1991 (Topping & Lindsay, 1992a) and two large-scale studies (Lloyd et al., 2015; Tymms et al., 2011). These reports provide valuable information on the purported positive effects of Paired Oral Reading methods; however, critical limitations confound any definitive conclusions about the influence of these methods on student reading achievement outcomes.

Topping and Lindsay's (1992a) Literature Review

Topping and Lindsay's (1992a) literature review is an important publication in the Paired Oral Reading literature. It is the only known publication that attempts to synthesize and report the outcomes of a single method of Paired Oral Reading. The literature review entails two sections of synthesis: a review the literature from 1976-1991 and reported outcomes from the Kirklees Paired Reading Project (Topping, 1990). Of the 53 studies conducted from 1976-1991, only a minority (30%) of outcomes were published in peer reviewed journal articles (n = 16). The remaining 37 studies consisted of completed theses/dissertations (n = 10, 19%), conference reports of school projects completed by classroom teachers or school psychologists (n = 13, 25%), outcomes reported in published books (n = 3, 6%), or unpublished data (n = 11, 20%).

The Topping and Lindsay (1992a) review reported accelerated reading achievement for students who participated in Paired Oral Reading. Across all studies, 1976-1991 (N = 53), students (tutor and tutee, N = 1,012) who participated in Paired Oral

Reading achieved growth in word-reading accuracy 4.23 times greater than expected and growth in reading comprehension 5.37 times more than expected. Critically, this analysis omits comparisons between treatment and control groups; it merely reports the growth of all participants across all projects. The reported achievement growth was measured using a questionable calculation called a "*Mean Ratio Gains"* (see Limitations section, this chapter). This calculation compares student growth in achievement compared to expected growth in achievement as determined by the norms of the reading assessment used. Although intended to be interpreted as a type of effect size (Topping, 1990), the Mean Ratio Gains is actually an indicator of raw growth rather than an effect size and should be interpreted with caution.

Topping and Lindsay (1992a) also reported effect sizes for a sub-sample of comparison studies present in the literature (n = 12). This analysis combined studies using parent and peer tutoring and includes the growth of peer tutors compared to readers serving in a control group. Overall, results indicated large effect sizes in accuracy (Δ = 2.12, SD = 2.26) and comprehension (Δ = 1.63, SD = 1.33) for students who participated in Paired Oral Reading. Again, these results should be interpreted with caution because the combined sample size (N = 230) was likely underpowered for the number of studies (n = 12) used in the analysis, as indicated by the erratically reported standard deviations (Maxwell, 2004; Vadillo et al., 2016). Indeed, the lower limits of both reported standard deviations approach null, or even negative outcomes.

As the only known systematic review of a Paired Oral Reading method, Topping and Lindsay's (1992a) report is a significant contribution to the Paired Oral Reading

literature. However, the pervasive use of small sample sizes ($\mu = 19.1$) and sporadic use of control groups ($n = 12$) in the surveyed literature ($N = 53$) limit any interpretation of this analysis. These authors (Topping & Lindsay, 1992a) also summarized findings from the Kirklees Paired Reading Project reported by Topping in 1990. Because the literature review only summarized major findings, the Kirklees Paired Reading Project will be considered in its own section, as reported by Topping (1990).

Kirklees Paired Reading Project

The Kirklees Paired Reading Project was a series of Paired Reading tutoring projects undertaken by various primary and secondary schools within the Kirklees Educational Authority (comparable to a U.S. school district) between 1984-1987. Essentially, the analysis completed by Topping (1990) is a synthesis of school-reported projects ($n = 155$), rather than empirical study conducted by researchers and should be interpreted as such. Similar to Topping and Lindsay's (1992a), this report provides a Mean Ratio Gain for all Paired Oral Reading participants across all projects and includes a subanalysis that reports an effect size for projects using control groups. Across all the Kirklees Paired Reading projects, students ($N = 2372$) who participated in Paired Oral Reading achieved a mean ratio gain 3.27 times greater than expected in accuracy and 4.39 times greater than expected in reading comprehension.

Again, limited Kirklees projects ($n = 34$) utilized control groups to compare student achievement in Paired Oral Reading. Using these studies, Topping (1990) reported large effects for reading accuracy ($\Delta = 0.87$, $SD = 1.04$) and reading comprehension ($\Delta = 0.77$, $SD = 0.72$). Similar to Topping and Lindsay's (1992a)

literature review, the standard deviations for these effects were quite large, approaching null or negative outcomes on the lower limit, and thus, should be interpreted with caution. Sample sizes are not reported for these studies; however, the larger Kirklees data set indicates a small average sample size ($\mu = 15.3$) across all studies. Further, of the 34 projects that used control groups, only 12 reported enough data for effect sizes to be calculated in reading comprehension.

The Kirklees Paired Reading Project is an important contribution to the Paired Oral Reading literature. Although cited for its attention to authentic learning environments, the limitations in sample size and lack of control groups in this analysis mirror those found in Topping and Lindsay's (1992a) literature review and reflect the larger limitations of the then-extant Paired Reading literature. Although the reported "literature on paired reading is substantial" (Topping & Lindsay, 1992a, p. 199), the lack of rigor present in the data severely limits any meaningful interpretation.

Fife Peer Learning Project

Another large-scale implementation of Paired Reading was the Fife Peer Learning Project, described in various reports (Topping et al., 2011, 2012; Tymms et al., 2011). The project occurred across 15 months at 143 elementary schools in Scotland. Twelve different conditions were randomly assigned to schools, with the conditions using combinations of same-age tutoring, cross-age tutoring, once-per-week tutoring, thrice-per-week tutoring, math-only tutoring, reading-only tutoring, and math/reading-combined tutoring. All tutoring sessions lasted 30 minutes.

Reading outcomes for the combined study were measured using the Performance

Indicators in Primary Schools (PIPS; Tymms, 1999), before and after the intervention. Notably, the baseline data for this report were collected nine months before the beginning of the 15-month intervention. Post-intervention data collection occurred within the same month that the intervention concluded. An overall effect size of 0.22 was reported for reading comprehension in conditions that used Paired Reading. h cross-age tutoring appearing more effective than same-age tutoring and no statistical difference between once weekly and thrice weekly tutoring. Unfortunately, the Fife Peer Learning Project report is opaque. Tutor and tutee outcomes were aggregated and no reports of specific subgroup outcomes (e.g., for thrice-weekly same-age tutee), are reported. Analysis of these data could have provided valuable information about the role of dosage in Paired Oral Reading and the specific outcomes for peer tutors.

Report by Education Endowment Foundation

One final executive summary reports a large-scale trial (N = 2,736) of paired reading conducted in the U.K. by the Education Endowment Foundation (EEF; Lloyd et al., 2015). This study used Year 9 (Grade 8 U.S.) pupils to peer tutor Year 7 (Grade 6 U.S.) for 20 minutes a week for 16 weeks. Randomization between treatment and control groups occurred at the classroom level (N = 160) and all tutors and tutees were assessed prior to the intervention using the New Group Reading Test (NGRT; Burge et al., 2010). Tutoring sessions used the paired reading method (R. Morgan, 1976) 20 minutes per week, for 15 weeks, before students were reassessed on the NGRT. The analysis indicated that paired reading had no influence on the reading outcomes of reading tutees (-0.28, p = 0.672) or tutors (-0.91, p = 0.125), and the authors did not recommend Paired

Reading as an effective practice in secondary schools.

The EEF study had several merits, such as randomization at the class level, power analysis to predict adequate sample size, and multilevel analysis. However, the project lacked a theoretical framework and an adequate review of the literature, with critical implications. Primarily, the authors neglected any consideration of oral reading proficiency in their analysis. Paired Oral Reading methods are generally seen as oral reading fluency interventions (Kuhn & Stahl, 2003; Rasinksi & Hoffman, 2003), and this oversight influences the design and outcomes of the study. Reported data indicate that 88% of student tutees were at or above expected levels of reading comprehension proficiency prior to intervention. Notwithstanding, all students in the treatment group participated as a tutee, regardless of initial reading proficiency. Given the contributive role of reading fluency to overall comprehension (Sabatini et al., 2019), it is likely that a vast majority of the sample who participated in this oral reading fluency intervention did not need an oral reading fluency intervention.

Further, the reported dosage of the intervention was exceptionally low. Students only participated in paired reading for a total of 8 hours over 16 weeks. The actual time spent reading by tutors and tutees was likely even lower; the dosage included pre-reading steps such as selecting a book and post-reading steps such as completing logbooks. The reported null results from this study are not unsurprising, given the high preexisting levels of reading achievement and low dosage.

Summary of Literature Reviews and Large-Scale Studies

R. Morgan's (1976) method of Paired Oral Reading has been investigated using school-level project reports, a literature review, and two large-scale studies. However, critical flaws hamper any solid conclusions from data regarding the efficacy of Paired Oral Reading. Early attempts, such as Topping's (1990) report of the Kirklees Paired Reading Project and Topping and Lindsay's (1992a) literature review suffered from low sample size, extremely limited use of control groups, and an over-reliance of school-reported projects. Suffering other design issues, results from the Fife Paired Learning Project (Tymms et al., 2011) indicated a 9-month gap between baseline scores and the beginning of the intervention. Further, for a project explicitly designed to determine the influence of peer tutoring and dosage on reading outcomes (c.f. Tymms et al., 2011, p. 271), it is curious that results did not adequately report differences on either of these outcomes. The most recent large-scale trial conducted by the Educational Endowment Foundation (Lloyd et al., 2015), studied a sample that was likely already proficient in oral reading fluency. These flaws complicate drawing definitive conclusions from the Paired Reading technique.

Synthesis of Tutee Fluency and Comprehension Outcomes

The Paired Oral Reading literature also consists of quasi-experimental and single group design studies. The following synthesis explicates the fluency and comprehension outcomes from the extant literature, including reports previously mentioned. The fluency and comprehension outcomes of the Paired Oral Reading literature are generally positive;

however, much remains unknown about the specific effects of Paired Oral Reading and important moderating factors, on the outcomes of student reading achievement.

Fluency Outcomes in the Paired Oral Reading Literature for Tutees

The relevant literature suggests that students who engage in Paired Oral Reading experience positive reading fluency outcomes. Positive results are reported across several measures including oral reading rate (Brown et al., 2018; Flood et al., 2005), silent reading rate (Flood et al., 2005), oral reading accuracy (Joscelyne, 1991; Limbrick et al., 1985), oral reading prosody (Brown et al., 2018; Young et al., 2015), and word recognition (Cook et al., 1980; McAllister, 1989; Young, Pearce, et al., 2018). The following sections will report the fluency outcomes of Paired Oral Reading, considering aspects of rate, accuracy, and prosody (Kuhn et al., 2010) in turn.

Rate. Reading rate is most often measured as the number of words read correctly in one minute. Students can be timed for exactly one minute to calculate rate; however, in some studies the students read a complete text from which an average words-correct-per-minute score was calculated (WCPM; cf., A. Morgan et al., 2000; Young, Pearce, et al., 2018). Studies in Paired Oral Reading have reported accelerated growth in reading rate (Brown et al., 2018; Fiala & Sheridan, 2003; Hermsmeyer, 1999; Thornton, 2012; Vo, 2011), in some cases reporting high effect sizes (d = 1.08; Young, Pearce, et al., 2018). Alternately, other studies have suggested that Paired Oral Reading did not promote oral reading rate when compared to a control group (Downs et al., 2020; Strong & Traynalis-Yurek, 1983).

The relationship between dosage of Paired Oral Reading and rate gains remains unclear. Two studies reported nearly identical growth in oral reading rates (21.6 WCPM; Downs et al., 2020; 20.7 WCPM; Young et al., 2015) with students in similar grades. However, the dosage of Paired Oral Reading differed significantly between these studies; Young et al. achieved this outcome with only 400 minutes of intervention, whereas students in Downs et al. read orally for approximately 900 minutes. Brown et al. (2018) and Flood et al. (2005) also present similar outcomes with a major discrepancy in dosage. The tutees in Brown et al. experienced an increase of 15.9 WCPM after 1,425 minutes of intervention, whereas Flood et al. (2005) report a 15.3 word per minute increase with only 200 minutes of intervention. Tutor type might help account for these differences. The studies with accelerated outcomes—Young et al., (2015) and Flood et al.—used adult tutors whereas Brown et al. and Downs et al. used same-age peer tutors. Further research is needed to examine the role of dosage on reading rate outcomes associated with paired reading practices and the potential influence of adult and peer tutors.

Accuracy. The literature on Paired Oral Reading also reports mixed outcomes for students in word reading accuracy. Several studies have indicated positive outcomes in word reading accuracy. Bush (1983) reported 12 months of growth in reading accuracy in only five months and Joscelyne (1991) reported five months of accuracy growth in a single month after using Paired Oral Reading. Young, Pearce, et al. (2018) reported that accuracy growth for tutees in Paired Oral Reading nearly doubled that of the control group. Several smaller studies with small sample sizes also reported growth in word reading accuracy (Gautrey, 1988; Lees, 1985; Low & Davies, 1988; O'Hara, 1985;

Richardson, 1986).

Other studies report null outcomes for accuracy among students using Paired Oral Reading. Two similarly designed studies implemented in third-grade classrooms (Brown et al., 2018; Downs et al., 2020) reported more growth in word reading accuracy from controls than from students who used Paired Oral Reading. Several other studies have also indicated little or no growth in accuracy as a result of Paired Oral Reading (Gately, 2014; Law & Kratochwill, 1993; B. V. Miller & Kratochwill, 1996). B. V. Miller and Kartochwill reported no statistically significant growth in accuracy for tutees, but also noted that tutees who received a higher dosage of Paired Oral Reading achieved higher levels of accuracy. Essentially, the influence of Paired Oral Reading on word reading accuracy—a targeted outcome—remains unclear.

Prosody. Measures of prosody are a recent addition to the Paired Oral Reading literature, appearing in only four studies (Brown et al., 2018; Young et al., 2015, Young, Pearce, et al., 2018, Young et al., 2020). These studies utilized the Multi-Dimensional Fluency Scale (MDFS; Zutell & Rasinski, 1991), and all indicated positive prosody outcomes for tutees. Two studies calculated effect sizes for tutee prosody gains, with each reporting large effects. (d = 1.44, Young et al., 2015; d = 1.03, Young, Pearce, et al., 2018). Further, raw score gains appear positive for tutees on the 16-point scale in the MFDS (2.79; Young et al. 2015; 2.8, Young, Pearce, et al., 2018; 4.16, Young, Pearce, et al., 2018; 4.4, Brown et al., 2020; 5, Young et al., 2020). These outcomes may be influenced by tutee grade level; the study with the smallest reported gains (Young et al., 2015) used a sample composed primarily of Grade 4 and 5 students, whereas the other

three studies used Grade 2 and Grade 3 students. These preliminary results demonstrate positive outcomes in prosody for tutees who read aloud synchronously with partners; however, more research is needed to confirm these findings.

Comprehension Outcomes in the Paired Oral Reading Literature

Studies in the Paired Oral Reading research corpus have measured reading comprehension using grade- and age-equivalent assessments (Brown et al., 2018; Downs et al., 2020; McAllister, 1989); norm-referenced tests (Eldredge & Quinn, 1988; Winter 1988); reading inventories (Flood et al., 2005; A. Morgan et al., 2000, Young, Pearce et al., 2018); and progress in online reading instruction platforms (Young et al., 2015; Young, Pearce, et al., 2018). Typically, these tests measured reading comprehension before and after the Paired Oral Reading intervention; however, two studies included an interim measure of reading comprehension (Brown et al., 2018; Downs et al., 2020). Comprehension outcomes tend to be positive across these measures; however, any differences between student grade level and tutor type remain unspecified.

Studies that measured reading comprehension using reader age, reading levels, or grade-level equivalency assessments have consistently reported positive comprehension growth associated with Paired Oral Reading. Often, reported achievement is greater than a single year's growth (Brown et al., 2018; Downs et al., 2020; McAllister, 1989; Sokal et al., 2005), and more rarely, two grade levels of growth (Cook et al., 1980; A. Morgan et al., 2000). Several studies concluded that below-proficient tutees achieved grade-level expectations by the end of the Paired Oral Reading intervention (Brown et al., 2018;

Eldredge & Quinn, 1988; McAllister, 1989).

Outside of the reviews and large-scale studies of Paired Reading, only four studies reported effect sizes for tutee comprehension outcomes (Almaguer, 2005; Downs et al., 2020; Young et al., 2015; Young, Pearce, et al., 2018). These effect sizes range from .35 (Downs et al., 2020) to 1.44 (Young et al., 2015), indicating medium to large comprehension effects from Paired Oral Reading practices. These reported effects are larger than the small effect reported for the Fife Peer Learning Project (*ES* = 0.22; Tymms et al., 2012), similar to the effects reported in the Kirklees Paired Reading Project (*ES* = 0.77; Topping, 1990), and smaller than the large effects reported Topping and Lindsay's (1992a) literature review (*ES* = 2.77). The variability among these comprehension outcomes could be viewed as a finding on its own. Clearly, the influence on Paired Oral Reading practices on comprehension and the related variables is not yet well understood and merit further scrutiny.

Indeed, several variables could potentially moderate comprehension outcomes for students who participate in Paired Oral Reading, most notably intervention dosage and tutor type. Currently, the literature presents several outcomes that appear to confound one another. For example, the largest reported effect size (*ES* = 2.77; Topping & Lindsay, 1992a) occurred across 12 studies averaging 50 sessions, whereas the smallest reported effect on comprehension (*ES* = 0.348; Downs et al., 2020) occurred in a single study across 95 sessions. Curiously, these reports represent similar tutee populations (*N* = 177, Downs et al., 2020; *N* = 190, Topping & Lindsay, 1992a). Further, Topping and Lindsay (1992a) reported larger mean ratio gains for comprehensions from projects involving peer

tutors rather than adult tutors, but more recently, studies with adult tutors have reported larger effect sizes than studies with peer tutors (*ES* = 0.348, Downs et al., 2020; *ES* = 1.44, Young et al., 2015; see also Almaguer, 2005; Young, Pearce, et al., 2018).

Interestingly, although Paired Oral Reading is generally intended to increase oral reading opportunities, increases in comprehension achievement appear to be the most frequently reported academic outcome in the Paired Oral Reading literature. However, the specific degree of growth remains unclear. The reported positive outcomes vary from minimal (Tymms et al., 2011) to especially large (Topping & Lindsay, 1992a). Further, the effect of moderating variables such as tutor type or dosage remains enigmatic. Additional analysis of Paired Oral Reading is needed to better determine the actual effects of Paired Oral Reading on comprehension outcomes.

Academic Outcomes for Peer Tutors

At its core, Paired Oral Reading is a tutoring procedure; a higher-level reader provides oral reading support for a lower-level reader via regular synchronous reading sessions. The academic outcomes of tutees are generally the most salient findings for researchers and practitioners while tutor outcomes are less frequently addressed. When adult tutors are used, issues of tutor gains of reading achievement are inconsequential. However, the use of peer tutors poses an important question, one unresolved in the extant literature: what is the influence of Paired Oral Reading on the academic outcomes of peer tutors of Paired Oral Reading? This question is complicated by how tutor success is measured; should tutor achievement gains be compared to tutee gains, expected gains, or similar achievement among control?

Studies have consistently reported that peer tutors experience academic gains similar to or greater than tutees (Brown et al., 2018; Downs et al., 2020; Jocelyne, 1991; Low et al., 1987; Topping, 1990; Topping & Bryce, 2004). Other studies have purported that peer tutors achieve greater gains than would be expected according to reading assessment norms (Topping, 1990; Topping & Whiteley, 1993). These findings may indicate that higher level readers experience accelerated reading achievement, but do not necessarily indicate that these readers receive additional benefit from Paired Oral Reading.

Heterogenous comparisons between peer tutor and tutee may be an inadequate measurement of the influence of Paired Oral Reading on tutor reading outcomes, given the differentiated gains in achievement by lower level and higher-level readers (Stanovich, 1986). A more effective contrast is comparing the achievement of tutors who participated in Paired Oral Reading to students in a control group who did not participate in Paired Oral Reading. Several studies that measure tutor achievement using this method indicate that achievement among tutors is similar to the achievement gains of control-group peers (Brown et al., 2018; Diaper, 1989; Downs et al., 2020; Joscelyne, 1991). In their role as higher level readers, peer tutors "might be viewed as the members who can facilitate [Paired Oral Reading] …but are not themselves the beneficiary of this reading strategy" (Downs et al., 2020, p. 130). Although peer tutors may not be "hindered in their reading development" (Brown et al., 2018, p. 550), the benefit they receive may be similar to the achievement derived from extended oral reading practice and not Paired Oral Reading itself. One exception to these findings comes from Topping and Lindsay

(1992a) who reported that peer tutors of Paired Oral Reading experienced large gains in reading accuracy ($\Delta = 1.71$, *SD* = 1.57) and reading comprehension ($\Delta = 1.24$, *SD* 0.61) compared to control readers. Importantly, Topping and Lindsay's (1992) finding was comprised of four studies with a combined sample of only 40 students.

Outcomes in Tutee Reading Attitudes

The relevant literature frequently relates Paired Oral Reading with improved reader attitudes. While many authors have mentioned this purported relationship anecdotally (Almaguer, 2005; Brown et al., 2018; Burdett, 1986; Flood et al., 2005; Henk, 1983; R. Morgan & Gavin, 1988), several studies have attempted to formally measure affective outcomes (Downs et al., 2020; Ellis, 1995; Hermsmeyer, 1999; Ottley 2003; Topping & Bryce, 2004; Young, Pearce, et al., 2018). Common reasons cited for improved reader attitudes include the opportunity for student self-selection of text (A. Miller et al., 1986), access to a wider range of texts (Klvacek, 2019; Topping & Lindsay, 1992b), improved comprehension (Henk, 1981), making reading more enjoyable (B. V. Miller & Kratochwill, 1996), social interaction/collaboration (Downs et al., 2020; Rasinski & Fredericks, 1991), and use of authentic, connected texts (Flood et al., 2005).

Despite the common contention that Paired Oral Reading activities promote positive reading experiences that influence reader attitudes, the relationship has not been determined definitively. The research has reported positive, null, and negative attitude outcomes for students who orally read in pairs across repeated sessions. The most common instrument used to measure reader attitudes in the literature has been the Elementary Reading Attitude Survey (ERAS; McKenna & Kear, 1990). This survey uses

images of the popular Garfield cartoon character on a 4-point Likert scale to measure student attitudes toward academic and recreational reading. This literature review found eight Paired Oral Reading studies that measured reader attitudes using the ERAS before and after a Paired Oral Reading intervention. Of these, three reported a statistically significant increase on the ERAS scale among participants (Hermsmeyer, 1999; Ottley 2003, Overett & Donald, 1998), two report a nonsignificant increase or stable attitudes (Sokal et al., 2005; Young, Pearce, et al., 2018), and two report a nonsignificant decrease in reader attitudes (Downs et a., 2020; A. L. Miller & Narrett, 1995). One study (Topping & Bryce, 2004) noted an increase in ERAS scores for many students but did not report any statistical results.

Tutee grade level may moderate these and other outcomes; more positive outcomes were reported in first grade (Hermsmeyer, 1999; Ottley, 2003) than second and third grade (Downs et al., 2020; A. L. Miller & Narrett, 1995). Intervention duration however does not appear to moderate these outcomes; these studies represent interventions that ranged from six weeks (Overett & Donald, 1998; Young, Pearce, et al., 2018) to 22 weeks (Sokal et al., 2005). Additionally, these studies include various implementations of Paired Oral Reading, such as home with a family member (A. L. Miller & Narrett, 1995), at home with a non-family adult tutor (Sokal et al., 2005), at school with an adult tutor (Young, Durham, et al., 2018), and at school with same-age peers (Downs et al., 2020).

Other studies have used difference scales to detect changes in reader attitudes. Weiss et al. (1989) found no changes in reader attitudes after a Paired Oral Reading

intervention using the Reading Attitude Scale (Heathington, 1975). Diaper (1989) used the Dundee Attitude Reading Test (DART; Ewing & Johnstone, 1981) to determine that overall attitudes declined for students who participated in Paired Oral Reading. However, a post hoc analysis of these results revealed that attitudes increased for students who began the study with negative or moderate attitudes toward reading. Tymms et al. (2011) report mixed outcomes for tutees who read orally in pairs with same-age and older-age peers using the Performance Indicators in Primary Schools (PIPS; Tymms, 1999) assessment.

Finally, several studies in the literature used non-standardized surveys or questionnaires to measure reader attitudes. The surveys or questionnaires were completed by the tutee and, in some cases, the parents regarding their perceptions of the tutee at the end of a Paired Oral Reading intervention. These studies have reported a largely positive influence of Paired Oral Reading on reading attitudes for tutees (Burdett 1986; Gautrey, 1988; Goudey, 2009; Law & Kratochwill, 1993; Topping, 1990); however, other studies indicated mixed or negative influences (DeAngelo et al., 1997; Gately, 2009).

Possible Reasons for Mixed Tutee Attitude Outcomes

The relevant literature offers several explanations for the various outcomes on reader attitudes associated with Paired Oral Reading interventions. These explanations include short intervention durations not influencing reading outcomes (Weiss et al., 1989; Young, Pearce, et al., 2018), possible lack of sensitivity or ceiling effects of the ERAS (Hermsmeyer 1999; Sokal et al., 2005), inaccuracies from survey/questionnaire response

data (Gately, 2014; Goudey, 2009), and attitudes toward Paired Oral Reading moderating general reading attitudes (Downs et al, 2020).

Limitations of the Paired Oral Reading Literature

As intimated by aforementioned incongruities, a variety of limitations exists in the extant literature on Paired Oral Reading. Many studies reported small sample sizes, likely underpowering any statistical analysis. At times, these studies used 10 or fewer students in a pre/post quasi experimental design (cf., J. Barrett, 1986; Leach, 1990; Lees 1985). Other issues occurred in two studies with sample sizes of 2,372 (Topping 1990) and 5,179 (Tymms et al., 2011). These large-scale implementation reports display inconsistencies in dosage, protocol, site details, and inadequate reporting (or complete omission) of control groups and subgroups.

Additionally, many Paired Oral Reading studies, particularly those completed in the U.K., merely describe paired reading projects implemented at individual schools, rather than reporting rigorous evaluation of the practices. These project reports lack critical design details such as frequency and length of Paired Oral Reading sessions or use of comparison groups (B. V. Miller & Kratochwill 1996). The Paired Oral Reading literature from the U.K. throughout the 1980s and 1990s is replete with these sampling and design issues, limiting interpretation of the results from these studies.

Another critique leveled at the Paired Reading literature is the persistent reporting of Mean Ratio Gain (Topping, 1990; Topping & Lindsay, 1992a). A mean ratio gain is student growth in reading age divided by chronological time. The result is a ratio that measures how much a student progressed in months compared to the time of the

intervention. This method emerged as a type of effect size to compare the reading progress of heterogeneous readers within a study to compare reader growth across studies (Topping, 1990). However, the Mean Ratio Gain calculation "is an approximation of extremely doubtful statistical validity" (Topping, 1990, p. 228) because the calculated effect does not account for the standard deviation across participants, a critical component of effect size calculation. Instead, reader progress is calculated in relation to chronological time (Topping & Lindsay, 1992a).

Many studies in the relevant literature investigated Paired Oral Reading interventions conducted in the tutee's home with a caregiver serving as tutor (Ellis 1995; Hermsmeyer, 1999; Howell, 2008; Lindsay et al., 1985; A. Miller et al., 1986; B. V. Miller & Kratochwill, 1996; Poliak, 1998). These studies lack measures of treatment fidelity, limiting findings from the reported research (Ellis 1995; Hermsmeyer, 1999; Howell, 2008; Lindsay et al., 1985; A. Miller et al., 1986). Further, studies that included measures of treatment fidelity report low levels (Goudey, 2009; Law & Kratochwill 1993; B. V. Miller & Kratochwill 1995; A. L. Miller & Narrett, 1995). A related limitation is the sample attrition in home Paired Oral Reading studies (Ellis, 1995; B. V. Miller & Kratochwill, 1995; Poliak, 1998). In these studies, a majority of participants who began did not finish the study, potentially biasing the reported results.

Another prevalent limitation of the Paired Reading literature is the repeated use of the Neale Analysis of Reading Ability assessment (NARA; Neale, 1958). The NARA is a reading assessment commonly used in the U.K. to calculate reading age. However, the assessment reportedly inflates the progress of reading achievement at lower levels; one

month of reading gain at the 6-year-old level represents a larger gain than one month of reading gain at the 8-year-old level (Topping, 1990). Thus, reading progress is potentially overestimated for younger or weaker readers, which is problematic for an intervention that investigates the academic progress of younger and lower-level readers. Other critiques leveled against the NARA concern inter-assessment reliability (Pumfrey, 1984; Stothard & Hulme, 1991). Reportedly, Form 1 and Form 2 of the NARA are not parallel, potentially influencing inaccurate estimates of growth between pre- and post-assessments. Form 1 of the NARA is purported to be a more complex assessment than Form 2 (Stothard & Hulme, 1991), though some dispute this claim (McKay, 1996). These issues with the NARA could have potentially limited findings from studies on Paired Oral Reading conducted in the U.K. Of the 51 studies cited in Topping and Lindsay's (1992a) literature review, 34 (67%) used the NARA assessment.

This field of study is also limited by diverse reporting from peer-tutor studies comparing treatment and control groups. One technique has been to compare the academic outcomes of the treatment tutees and/or tutors against a control group consisting of low, middle, and high-level readers (Diaper, 1989; Topping & Lindsay, 1992a). This practice is a heterogeneous comparison, evaluating lower or higher-level readers in the treatment group against the combined achievements of the control group, complicating any comparative analysis between treatment and control groups.

Another problematic technique is comparing control-group outcomes with combined peer-tutor and tutee outcomes (Topping, 1990; Tymms et al., 2011). As previously evidenced, the roles of tutor and tutee have fundamental differences; tutees

receive support from the tutors using books selected by or for the tutees. This simple, but critical, characteristic influences the achievement progress of the student in each role. Collapsing student outcomes into a single analysis fails to capture the clear differences between the tutor/tutee roles in Paired Oral Reading.

Perhaps the most apt comparison is measuring the outcomes of tutees against the academic progress of similarly lower-level readers in a control group while monitoring the progress of tutors against higher level readers in the control group (Brown et al., 2018; Downs et al., 2020; Eldredge & Quinn, 1988). In this case, the achievement of tutor and tutee cohorts is measured against similarly achieving peers, generating a more accurate analysis of the influence of Paired Oral Reading.

Conclusion

The variations of Paired Oral Reading have sought to improve the progress of remedial readers through direct tutelage with extended connected text. Although initially invoking disparate theoretical frameworks, current research frames for these practices use cognitive theories, most commonly LaBerge and Samuels' (1974) theory of automaticity. The extant literature generally reports positive fluency and comprehension outcomes for students who engage in Paired Oral Reading, at times within six weeks or less (Crombie & Low, 1986; Young et al., 2015). At the elementary level, these positive outcomes are reflected in studies with tutees in Kindergarten (Cadieux & Bourdrealt, 2005) through fifth grade (Winter, 1988; Young et al., 2015).

Despite consistently reported positive academic outcomes, drawing specific conclusions from the extant research is problematic. Characteristics such as tutor type,

setting, dosage, and tutee age vary across the literature, potentially influencing reported outcomes. Study quality and reporting inconsistencies further compound efforts to understand the specific influence of Paired Oral Reading on academic outcomes. Put simply, the aggregate effects on comprehension and fluency outcomes for students who engage in Paired Oral Reading are not yet well understood. Further, it is unknown to what degree factors such as setting, tutor type, dosage, and tutee age moderate the academic outcomes of Paired Oral Reading. The lack of synthesis across these iterations of Paired Oral Reading is vexing; the extant literature fails to elucidate the myriad variables of a highly touted instructional practice for improving reading achievement among weaker elementary readers. Ultimately, additional analyses are needed to clarify the academic outcomes of Paired Oral Reading and to identify significant factors that influence those outcomes.

CHAPTER 3

METHODS

Supporting myriad readers within a single classroom is a perennial challenge for classroom teachers. Pairing a struggling reader with a more proficient reader–whether adult or peer–is one common method that teachers use to support such readers. A review of relevant literature revealed several lines of research where mixed-proficiency pairs orally read extended text in unison. These studies generally report positive oral reading fluency and reading comprehension outcomes. Critically, effect of Paired Oral Reading on student reading achievement outcomes remains unanalyzed. In response, an analysis sought to determine the effects across reported literacy outcomes, investigate which study-level factors influence the outcomes, and calculate fluency and comprehension effect sizes for tutors and tutees while controlling for significant variables. The synthesized, and scrutinized, outcomes from this analysis informs an improved and nuanced understanding of Paired Oral Reading in support of literacy development.

Research Questions and Method

Using meta-analytic methods, this dissertation addressed the following research questions.

1. What is the effect of Paired Oral Reading on reading outcomes for tutees in grades K-6?
2. How do related factors (e.g., tutor type, dosage, year, publication type, and method used) moderate the reported outcomes?
3. What are the effects of Paired Oral Reading for tutees while accounting for significant moderator variables?

Meta-Analysis

Meta-analysis is a form of systematic review involving the "quantitative analysis and synthesis of a set of related empirical studies in a well-defined domain" (Bus et al., 2011, p. 270). Generally, meta-analytic methods follow a standard procedure. First, the researcher determines a "critical question worthy of review" (Alexander, 2020, p. 7). Next, the researcher commences an exhaustive search and screens studies for inclusion using predetermined criteria (Card, 2015). The researcher then records standardized mean difference outcomes from the included studies and codes imported study features (Pigott & Polanin, 2020). Finally, the reported outcomes are statistically analyzed to determine an aggregate effect (Borenstein et al., 2010). This dissertation study used meta-analytic techniques to determine the significance of reported academic outcomes of Paired Oral Reading and analyze the features as variables that influence those outcomes.

Screening and Inclusion Criteria

The reported literature review included a search of five academic databases and a timeframe of 1980-2020 using various terms related to Paired Oral Reading. The titles, abstracts and related details from this search were exported to inform the literature review. One finding from the literature review was that two methods, Paired Reading and Neurological Impress Method originated prior to 1980. To further identify reported outcomes, an identical search and exportation process was consequently conducted covering 1960-1980 and then merged with the existing 1980-2020 files. Although completed consecutively, these efforts comprise a unified search and the following describes the combined process.

As part of the literature review, the databases Academic Search Ultimate, APA PsychInfo, ERIC, OpenDissertations, and Professional Development Collection searched for studies that used variations of Paired Oral Reading from the years 1980-2020, and subsequently 1960-1980. The search included various combinations the following terms gleaned from reading resources related to the broader topic: *dyad read*, neurologic* impress method, "pair* read*," partner read*, "peer-assisted reading," "prime-o-tec," "reading-while listening," "listening while reading," "shadow read*," "shar* read*," "assisted reading," "duet read*," "duolog read*," paired assisted read*, paired partner read*, impress read*, read two impress, and "buddy read*."* From this initial search, titles and abstracts were reviewed for additional studies involving Paired Oral Reading. Further, reference lists from literature reviews involving Paired Oral Reading (cf. Kuhn & Stahl, 2003; Rasinski & Hoffman, 2003; Topping & Lindsay, 1992a) were reviewed to inform the search for relevant research.

All titles and abstracts were then reviewed for evidence of empirical research of Paired Oral Reading. Duplicate articles were removed, and the remaining studies were individually reviewed to determine if they met the following inclusionary criteria.

- Studies were published in a peer-reviewed journal, national conference proceeding, or a completed thesis/dissertation between the years 1960 and 2020.
- Participants included students in Grades K-6.
- Researchers used an experimental, quasi-experimental, or single-subject type design.
- Independent variables consisted of practices that used a higher level reader reading a text in unison with a lower level reader during regularly scheduled sessions within a school environment.

- Researchers measured fluency or comprehension as dependent variables.
- Studies occurred in schools where English was the primary language.

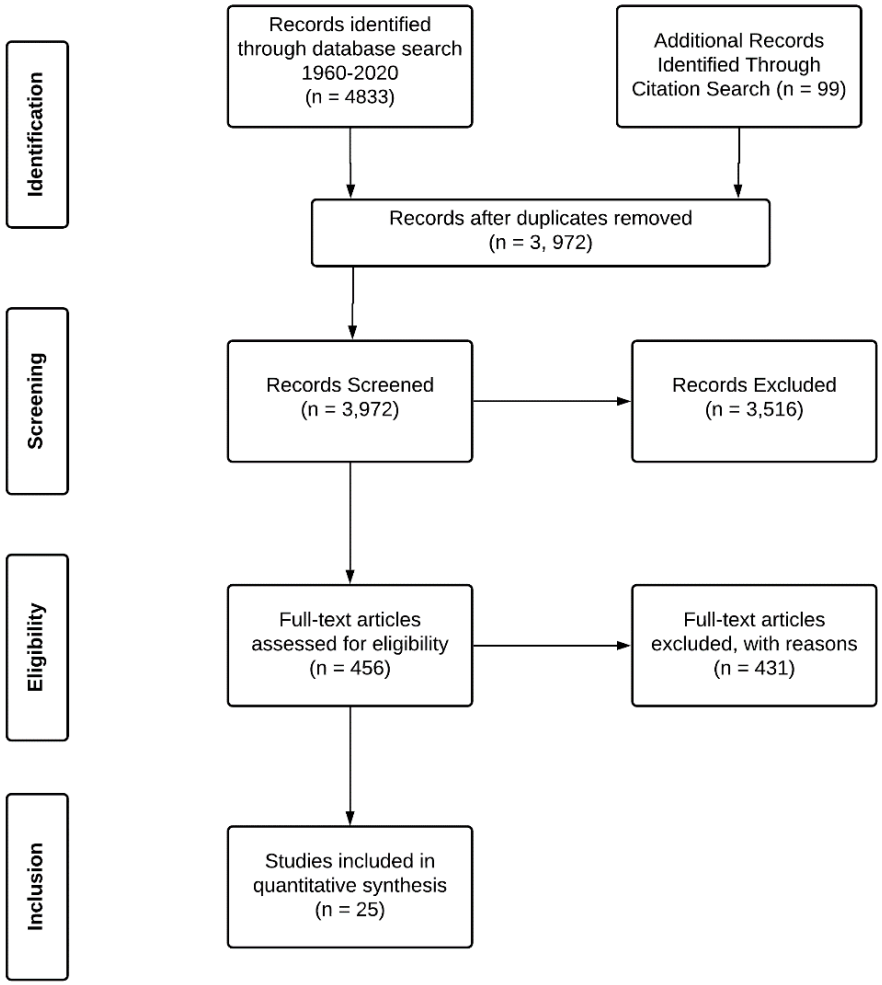
Exclusionary Criteria

This search process returned a high volume of references (N = 4,833). Many of these were not relevant to Paired Oral Reading or did not contain quantitative analysis of Paired Oral Reading outcomes. As indicated in Figure 3.1, all references were systematically sifted using the inclusion criteria until a final corpus for meta-analysis was identified. Studies lacking sufficient quantitative information to calculate effect sizes for tutees were excluded. Further, studies reporting duplicate data or interventions occurring in a non-school environment were also excluded. The studies identified during this process underwent a coding process followed by meta-analysis.

This meta-analysis included a coding scheme to capture important features of studies that met the inclusion criteria. Coding eligible studies is an essential component of meta-analytic methods because coding accounts for the research elements targeted in the analysis and can control for study artifacts through moderator analysis (Card, 2015). The coding process recorded study characteristics into a single Excel spreadsheet following an established coding manual. The scheme noted publication type (i.e., peer-reviewed journal, dissertation, conference report, or other), sample size, age of participants, tutoring method (i.e., Paired Reading, Neurological Impress Method, Dyad Reading, Read Two Impress, or other), tutor type (i.e., same age tutor, adult/cross age tutor), dosage (i.e., total minutes of intervention), duration (i.e., total weeks of intervention), study quality, fluency effect sizes, and comprehension effect sizes. The

Figure 3.1

Search and Screening Process, 1960-2020



resulting codes sets were then analyzed to calculate aggregate effect sizes and determine potential moderating variables (Pigott & Polanin, 2020). The coding manual used in this study is reported in the Appendix and further explained in Chapter 4.

Calculating Effect Sizes

As part of the coding process, a Hedge's g effect size (Hedges & Olkin, 1985) was calculated for each reported fluency and comprehension outcome. As shown in

Equation 1, Hedge's *g* is calculated by subtracting the posttest mean from the pretest mean and dividing by the pooled estimate of the population's standard deviation. Hedge's *g* is appropriate for this analysis because it calculates the standardized mean difference for longitudinal designs and is considered less biased than Cohen's *d* for small study samples (Card, 2015).

$$g = \frac{\sqrt{M_1 - M_2}}{s_{pooled}} \quad (1)$$

Two separate pools of Hedge's *g* effect sizes were calculated from studies meeting the inclusion criteria. The first pool included the effects that utilized treatment/control study designs, specifically excluding single group designs. This pool compared the growth of the treatment group with the growth from the control group. The second pool included all effects the met the inclusion criteria and measured the effect of treatment group outcomes. The first pool of treatment/control outcomes informed the majority of analysis in the study. The second pool of treatment-only outcomes was used in a single analysis as a sub-component of Research Question #1.

Data Analysis

Multilevel Meta-analysis

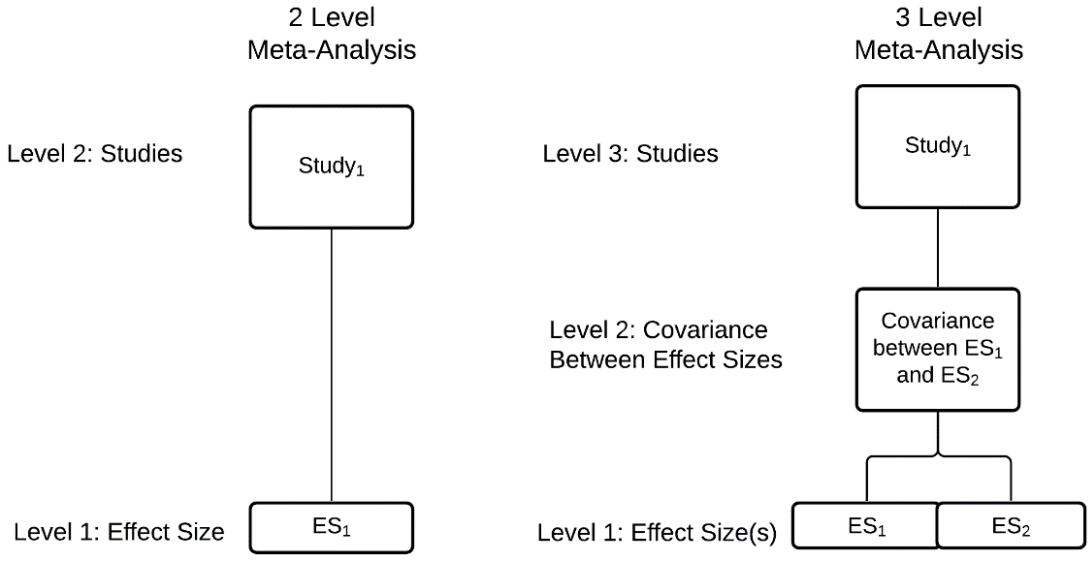
Meta-analytic approaches calculate aggregate effect sizes using outcome level and study-level data (Card, 2015). Typically, outcome effect sizes are nested within studies and covariances between studies are modeled within the analysis. The hierarchical structure reflected in meta-analytic designs indicates that all meta-analyses are inherently multilevel (Fernández-Castilla et al., 2020; Hox et al., 2018). The classical form of meta-

analysis is a two-level analysis with the first level of analysis consisting of effect sizes and studies forming the second level. However, other hierarchical structures of data are possible, and indeed desirable (Cheung, 2014, 2019; Fernández-Castilla et al., 2020; Hox et al., 2018).

One alternative hierarchical structure is a three-level model. Typically, these models extend the classical two-level analysis by adding an intermediary level of regression (Fernández-Castilla et al., 2020). The intermediary level models the covariance between effect sizes, allowing multiple outcomes to be included from a single study, facilitating greater flexibility with the study design (Hox et al., 2018). Overall, the added sophistication of the three-level model promotes the robustness of the design and increases the accuracy of outcomes. Figure 3.2 displays the hierarchical structures extant in classical and multilevel meta-analysis.

Figure 3.2

Comparing Classical and Multilevel Meta-Analysis



Despite classical meta-analysis being technically a multilevel approach, multilevel meta-analysis commonly refers to meta-analytic designs that utilize three or more levels (Hox et al., 2018). The use of multilevel meta-analysis has grown rapidly since 2015, perhaps, in part, due to its ability to account for sophisticated hierarchies within a dataset (Fernández-Castilla et al., 2020). Including additional hierarchies through multi-level modeling can account for more covariance within the model, providing a more accurate analysis (Hox et al., 2018). One important implication is that multilevel meta-analyses model non-independent effect sizes particularly well, affording the inclusion of multiple treatment groups from a single study, or multiple outcomes within a single study (Cheung, 2019; Van den Noortgate et al., 2015).

The literature review reported in Chapter 2 revealed fluency and comprehension outcomes being most commonly reported in the eligible Paired Oral Reading research. Further, the research commonly reports multiple outcomes within a single study, and in some cases multiple treatment groups (e.g., Brown et al., 2018; Young, Pearce, et al., 2018). Multilevel analysis accounts for dependencies such as these by including additional levels of regression within the model, allowing greater flexibility in application. These and other affordances indicated that a multilevel analysis would be most appropriate study design for the analysis of Paired Oral Reading outcomes.

A three-level multivariate meta-analysis was employed to analyze the extant data. Specifically, level one modeled individual effect sizes, level two modeled covariance between effect sizes (within studies), and level three modeled covariance between studies (Fernández-Castilla et al., 2020; Hox et al., 2018; Van den Noortgate et al., 2015).

Equation 2 represents the proposed analysis:

$$d_{ij} = \gamma_{00} + r_j + u_{ij} + e_{ij} \quad (2)$$

In the equation, d_{ij} refers to the effect size of outcome i within study j. The remainder of the equation consists of a single fixed component and three random components. The fixed component γ_{00}, refers to the combined effect size and the random component e_{ij} refers to the random residual error term. The outcome term, fixed effect, and residual random term constitute the formula for a classical meta-analysis. The addition of two more random components provides the three-level structure. The random component u_{ij} models the variance of each observed effect size around its population mean while the random component r_j models the deviation of each study effect mean from the overall mean (Fernández-Castilla et al., 2020; Hox et al., 2018). The multilevel model can provide valuable information about the extant data, such as the overall effect of Paired Oral Reading, variability across studies, and the effect of potential moderators. All analyses were conducted using R (R Core Team, 2018), using the furniture (T. A. Barrett & Brignone, 2017), metafor (Viechtbauer, 2010), and dmetar (Harrer et al., 2019) packages.

Research Question #1

The first research question guiding this dissertation is *What is the effect of Paired Oral Reading on reading outcomes for tutees in grades K-6?* To address this question, a model representing Equation 2 determined the average reading outcomes for students who engaged in Paired Oral Reading. This analysis modeled paired synchronous oral reading outcomes as a fixed effect, with random effects added for between-outcome

variance and between-study variance (Fernández-Castilla et al., 2020), using the effect sizes from the treatment/comparison pool. This model allowed calculation of an average effect size and 95% confidence interval (Pigott & Polanin, 2020) for the pooled data. Importantly, no moderator variables were included in this initial model (Card, 2015); however, random effects were included to account for the hierarchical clustering. This model is essentially equivalent to the intercept-only null model recommended by Hox et al. (2018). A null model is desirable because it "provides a benchmark value of deviance…which can be used to compare models" (p. 43). Thus, this null model serves two important purposes within this analysis. First, the null model reports the weighted effect size of Paired Oral Reading across the treatment/comparison pool of data, while accounting for hierarchical clusters within individual studies and covariance across studies. Second, the null model operates as the comparison model for the meta-regression and multiple meta-regression performed in Research Question #2 and Research Question #3.

Three other similar models were also fit to answer Research Question #1. The first model performed the exact same analysis; however, it included the pool of treatment-only effect sizes, rather than the treatment/control effect sizes. The next two models added tutor type and reading outcome type as singular fixed effects with the intercept removed to model the outcomes of these factors. All models were fit using Restricted Maximum Likelihood (REML; Hox et al., 2018) to provide estimates of covariance among clusters and their significance.

In conjunction with calculating effect sizes, a test of heterogeneity among

outcomes was conducted (Del Re, 2015; Pigott & Polanin, 2020). A high degree of heterogeneity among outcomes suggests that study characteristics such as treatment duration or tutor type may influence the calculated effect size outcomes (Card, 2015; Del Re, 2015). Heterogeneity between effect size outcomes was measured using the I^2 statistic. This statistic measures the ratio of variability in effect sizes that is due to true differences among studies, reported as the percentage of unexplained variability in the model (Borenstein et al., 2017). A large I^2 statistic indicates the variability between studies is due to true differences between the studies, a small statistic indicates that variability is the result of sampling error (Del Re, 2015). I^2 is measured by subtracting the degrees of freedom from the total sum of squares and then dividing by the total sum of squares (Borenstein et al., 2015). Equation 3 presents the equation for this statistic:

$$I^2 = \left(\frac{Q-df}{Q}\right) \qquad (3)$$

Importantly, the I^2 statistic calculates the proportion of heterogeneity present in the study corpus (Borenstein et al., 2015). It does not measure the presence of outliers within the data or measure the influence of one particular outcome on the overall corpus. These characteristics were analyzed via a robust and innovative technique—Graphical Display of Study Heterogeneity (GOSH; Olkin et al., 2012). GOSH plots are a method of combinatorial meta-analysis whereby individual meta-analyses are run on all the possible $2^k{-}1$ combinations of effect sizes within the sample. By running combinations of the meta-analysis patterns, overly influential studies can be detected and potentially removed from the analysis (Olkin et al., 2012). This method is viewed as more sophisticated than the more typical leave-one-out method of heterogeneity detection and supports the

overall robustness of the meta-analysis.

The computational intensity of GOSH analysis increases radically as the number of analyzed effects increases (Olkin et al., 2012). This limited the current GOSH analysis in two ways. First, the GOSH analysis was fit using a two-level model that omitted covariance between-effects within studies. Second, the GOSH analysis was limited to one million random subsets of the data, instead of all possible combinations. Olkin et al. indicate that a random subset of analyses is an appropriate implementation of GOSH. In conjunction with the GOSH analysis, three supervised machine learning algorithms (Gaussian Mixture Models, Fraley & Raftery, 2002; k-Means Clustering, Hartigan & Wong, 1979; DBSCAN, Schubert et al., 2017) were used to identify distinct clusters within the GOSH plot. In this analysis, the I^2 and effect size data in the GOSH plot were standardized as z-scores. Each algorithm then analyzed the data to detect anomalies. These anomalies were subsequently refit to a new model and compared with the original fit to ascertain the better predictor of Paired Oral Reading outcomes. These issues were addressed using meta-regression moderator analysis in Research Question #2.

Research Question #2

The second question guiding this analysis was *How do related factors (e.g., tutor type, dosage, year, publication type, and method used) moderate the reported outcomes?* The Chapter 2 literature review indicates that important differences among outcomes, tutor type, dosage, and other variables are present in the extant literature, suggesting the presence of heterogeneity among outcomes in the extant data. Thus, a moderator analysis was used to determine the factors that potentially influenced heterogeneous outcomes

(Card, 2015). Table 3.1 lists and describes the moderating variables analyzed in this study.

Table 3.1

Description of Moderator Analysis Variables

Variable	Description
Year	Year of publication
Publication type	Type of publication, coded as peer-reviewed journal, thesis/dissertation, conference paper/presentation
Study design	Design of study coded as treatment-group only design, treatment/control design with nonrandomization, treatment/control design with randomization
Age	Age of tutee, as indicated in publication
Tutoring method	Method of tutoring received by tutee. Coded as Neurological Impress Method, Paired Reading, Dyad Reading, and Read Two Impress
Tutor type	Type of tutor as indicated in publication. Coded as adult, cross-age, or peer
Dose	Total dose of intervention received by tutee, calculated in hours
Duration	Total duration of tutoring, calculated in weeks
Sessions	Total number of sessions
Reading outcome	Type of reading outcome assessment used, coded as fluency or comprehension

Meta-regression is one form of moderator analysis that addresses how effect sizes vary across variables collected as coded (Deeks et al., 2021). Meta-regression operated within a multilevel framework allows a variable to be added to the null model as a fixed factor. This meta-regressed model calculated the effect size of the respective outcomes while accounting for the fixed factor. The meta-regressed model was then compared to the null model using the Loglikelihood Ratio Test (LRT: Hox et al., 2018) using statistical significance set at $\alpha = 0.05$. The moderators that were investigates using this meta-regression technique included year of publication, publication type, study design, tutee age, tutoring method, tutor type, total hours of intervention, total weeks of

intervention, and reading outcome type coded as fluency or comprehension. Any variable included in any meta-regressed model that accounted for significantly more variance than the null model was accepted as a significant moderator.

Research Question #3

The third research question that guided this study was *What are the effects of Paired Oral Reading for tutees while accounting for significant moderator variables?* Research Question #3 was answered using multiple meta-regression techniques. An extension of meta-regression, multiple meta-regression adds two or more fixed factors to the null model (Harrer et al., 2019). This procedure is advantageous because it predicts the overall effect size of outcomes while accounting for variables of interest. Multiple meta-regression typically accounts for greater amounts of variance, producing a more accurate estimate of effect (Pigott & Polanin; 2020; Tipton et al., 2019b). This analysis used an a priori stepwise model building (Fernández-Castilla et al., 2020; Hox et al., 2018) to conduct the multiple meta-regression. First, tutor type (adult/cross-age/peer) and reading outcome type (fluency/comprehension) were introduced to the null model as fixed effects. This model was then compared to the null model using the Logliklihood Ratio Test (Hox et al., 2018). The second multiple meta-regressed model retained the tutor type and outcome type fixed effects from the previous model and included all other significant moderators from Research Question #2 as fixed effects. This model was then compared to the first multiple meta-regressed model and the null model. Essentially, this final model reported the effect size of paired synchronous oral reading for tutees depending on tutor type across fluency and comprehension outcomes, while controlling

for significant moderating variables.

Conclusion

This dissertation study analyzed across studies the reading outcomes for students who engage in Paired Oral Reading. Via a thorough search and screening process, quantitative studies published in the last six decades were identified for inclusion in the analysis. All studies meeting the inclusion criteria were coded for subsequent analysis.

The multilevel meta-analytic and meta-regressive techniques used in this study afford a thorough and nuanced analysis of the extant literature. First, an aggregate effect size for reading outcomes in Paired Oral Reading studies was estimated using a null multilevel model that accounted for hierarchal clustering extant within studies. Heterogeneity among outcomes was then tested to detect whether study-level differences contributed to effect variance. Next, individual moderators were meta-regressed as fixed effects and compared to the null model to ascertain their effect on the reported outcomes. Subsequently, a stepwise multiple meta-regression introduced tutor type and reading outcome type to the model, calculating different outcomes and their significance for tutors and tutees across fluency and comprehension outcomes. Finally, a second block containing all statistically significant moderators was introduced to the multiple meta-regression, representing a final model that revealed tutee outcomes based on tutor type across fluency and comprehension while controlling for statistically significant variables.

This multilevel meta-analysis reflected many current best practices in meta-analytic methods, such as articulating an answerable question that contributes to the literature (Alexander 2020), appropriately accounting for dependent effects between

outcomes and between studies through a three-level analysis (Fernández-Castilla et al., 2020; Pigott & Polanin, 2020), inclusion of all relevant effect sizes in a single model (Tipton et al., 2019a), pre-planning specific analyses and limiting exploratory analysis (Tipton et al., 2019a), and investigating effect size variability via moderator analyses and meta-regression (Pigott & Polanin, 2020).

Further, the design of this meta-analysis includes many robust and innovative techniques that allow the research questions to be answered with enhanced accuracy and nuance. The three-level structure modeled covariance within a single study as part of the overall data structure. This design allowed greater flexibility with the analysis by including multiple treatment groups and multiple outcomes within a single study. Further, this structure allowed for fluency and comprehension outcomes to be included within a single model to parse out the potential differential effects identified in the literature review. The analysis was also enhanced by the meta-regression and multiple meta-regression techniques, which were used to produce more accurate estimates of Paired Oral Reading outcomes. As delineated in Chapter 4 and Chapter 5, this sophisticated meta-analysis of Paired Oral Reading methods facilitates a greater understanding of the extant literature, fluency and comprehension outcomes, and the relative influence of the coded variables. These outcomes provide a multi-faceted view of Paired Oral Reading techniques, with implications for future research and practice.

CHAPTER 4

RESULTS

Several methods of tutoring use a higher-level reader to chorally read connected text aloud with a lower-level reader. Termed Paired Oral Reading, researchers and practitioners alike have utilized such techniques to support developing readers. Despite nearly six decades of usage, minimal research has synthesized or attempted to calculate the effectiveness of these methods. This dissertation sought to understand the average effect of these methods on student reading achievement via meta-analysis. Five academic data bases were searched from the years 1960-2020 and systematically screened using an established inclusion protocol.

Study Coding

Prior to coding studies meeting the inclusion criteria, two researchers met to review the codebook and the existing codes generated from the literature review. Discussion topics during this first meeting included recording data from studies by individual effect, using a long format to record coded metrics in a spreadsheet, determining levels for evaluating study quality, and reviewing the definition of each code specifically to promote a shared understanding of each construct. The researchers concluded by coding a single study together to familiarize themselves with the coding process. Minor adjustments to the codebook were made after this initial session to facilitate efficient data collection. After the initial session, the coders independently coded four studies randomly selected from the corpus to compare interrater reliability.

The interrater reliability was 97.1% for codes applied to the four studies.

After the initial meeting to review the viability of the codebook and calculating the interrater reliability, the researchers met again to discuss variations across the studies, to review discrepancies, and evaluate the validity of the codebook. The majority of discrepancies during the initial coding cycle were minor clerical errors regarding the coding of tutor method, study quality, and calculating descriptive statistics from studies containing raw data. Further discussion during this second session included determining a protocol for resolving discrepancies and reviewing the long data format for individual effects. After this session, each study was coded twice, once by each researcher. The researchers met a third time to review the codes and discuss issues with a few remaining studies. For example, some studies reported weeks and total number of sessions, but not the number of minutes per session, other studies did not adequately report tutee data. These issues were resolved by consensual coding between the researchers and all studies were cross-checked for coding accuracy and consistency. Interrater reliability was not calculated for this process; however, accuracy between coders remained high and all discrepancies between codes were resolved. The researchers met for a total of nine hours during the coding process.

Description of the Study Sample

The coding procedure yielded a total of 87 effect sizes across 25 eligible studies. Of the 87 effects, 21 (24%) were derived from treatment-only study designs while 66 (76%) effects were reported in studies using treatment/comparison designs. The remainder of this section will describe various aspects of the sample contrived from the

treatment/comparison designs, as these designs comprised the majority of the meta-analysis. Notably, due to the inclusion of multiple effects from a single study in the meta-analysis, these descriptive statistics are reported across all effects rather than all studies.

The average sample size of among treatment/control designs was limited. The average sample size for reading tutees was 20.2 (*SD* = 17.2) and 16.8 (*SD* = 6.9) for control groups. The total sample of tutees (*n* = 1,333) and control students (*n* = 1,111) was similar. The combined sample size of all tutees across all treatment/control designed studies was 2,444. Table 4.1 provides the descriptive statistics for tutees in treatment/control designed studies.

Table 4.1

Descriptive Statistics for Reported Tutee Samples, Treatment/Control Effects

Variable	*M*	*SD*	Range	Total
Control Sample	16.8	6.9	8 - 40	1,111
Tutee Sample	20.2	17.2	6–82	1,333
Combined Sample	37.3	21.3	15-102	2,444

The continuous variables included in the analysis was publication year, tutee age, dose of intervention in hours, duration of intervention in weeks, and the total number of intervention sessions. Although the search range extended to 1960, there were no treatment/control effects identified prior to 1980. The average tutee age among all effects was 8.9 years (*SD* = 0.9). The small standard deviation indicates that Paired Oral Reading research has targeted a narrow age range. The three-time variables, dose in hours, duration in weeks, and number of sessions, exhibit a positively skewed distribution. In

other words, the upper-limit of these variables extend farther from the mean than the lower-limit of these variables. For example, the average dose of interventions among effect sizes was 11.3 hours (*SD* = 6.5), however the range extended as high as 23.75 hours. Similarly, the average duration of intervention effects was 10.2 weeks (*SD* = 6.1). how the range of effects extends 36 weeks. The average total number of sessions among effects was 43.1 (*SD* = 30.2) and the range extended to 180 sessions. The research that reported a high dosage of intervention consists of recent studies from the Dyad Reading (Brown et al., 2018; Downs et al., 2020) line of research. Table 4.2 reports the descriptive statistics for continuous variables among treatment/control design effects.

Table 4.2

Descriptive Statistics for Continuous Variables, Treatment/Control Effects

Variable	*M*	*SD*	Range
Publication year	2000.1	14.6	1980 - 2020
Tutee age	8.9	0.9	7.35 - 10.25
Dose in hours	11.3	6.5	5 - 23.75
Duration in weeks	10.2	6.1	4 - 36
Total sessions	43.1	30.2	18 - 180

The categorical variables used in the meta-analysis consisted of publication type, study design, tutoring method, tutor type, and reading outcome type. The majority (52%) of effects (*n* = 34) resulted from the Paired Reading literature, followed by Dyad Reading (*n* = 16; 24%), Read Two Impress (*n* = 8; 12%), and Neurological Impress Method (*n* = 8, 12%). The literature review and study screening did not locate any effects not using one of these four methods.

Three tutoring arrangements were located in the literature review: same-age tutors, adult tutors, and cross-age tutors. Same-age tutors constituted the majority (62%) of the sample (n = 40), followed by adult tutors (n = 19; 29%) and cross-age tutors (n = 7; 11%). Effects were coded as either a comprehension or fluency achievement outcome. Effects were split in half between comprehension (n = 33; 50%) and fluency (n = 33; 50%). The study designs used were nearly even between non-random assignment (n = 35; 53% and random assignment (n = 31; 47%). Table 4.3 lists the frequency of categorical variables used in the meta-analysis.

Table 4.3

Descriptive Statistics for Categorical Variables, Treatment/Control Effects

Variables	Frequency (n)	Percentage (%)
Publication type		
Peer reviewed journal	33	50
Thesis/dissertation	30	45.5
Conference presentation	3	4.5
Study design		
Random assignment	35	53
Non-random assignment	31	47
Tutoring method		
Neurological impress method	8	12.1
Paired reading	34	51.5
Dyad	16	24.4
Read Two Impress	8	12.1
Tutor type		
Adult	19	28.8
Cross-age	7	10.6
Same age	40	60.6
Reading outcome type		
Fluency	33	50
Comprehension	33	50

The average year of publication for treatment/control designed effects was 2000 (*SD* = 14.6). Again, although the search of literature extended to 1960, no treatment/control studies were located prior to 1980. Among specific Paired Oral Reading methods, Neurologic Impress Method had the largest range of research, with studies reported from 1980 to 2018. Read Two Impress had the smallest range, with treatment/control effects existing between 2015 to 2018. Table 4.4 reports the descriptive statistics for the year of publication for treatment/control designs by tutoring method.

Table 4.4

Method Descriptive Statistics by Year of Publication, Treatment/Control Effects

Method	Mean	*SD*	Range
Neurologic impress method	1999	14.3	1980 - 2018
Paired reading	1989	1.80	1986 - 1991
Dyad reading	2015	8.6	1988 - 2020
Read Two Impress	2017	1.55	2015 - 2018

Summary of Descriptive Statistics

The descriptive statistics resultant from the coding process supports the trends identified in the literature review; four methods of Paired Oral Reading have been researched during the past 60 years. As indicated by the dispersed span of publication for each method, the popularity of these methods occurred in waves, reflecting the aforementioned methodological and theoretical evolutions of these methods. Moreover, the data reveal wide variations in time variables and sample sizes among the four methods. Table 4.5 lists the selected treatment/control effects used in the meta-analysis in alphabetical order with relevant study details.

Table 4.5

Description of Effects Used in the Treatment/Comparison Meta-Analysis

Effect Number	Effect Name	Method	Tutor Type	Publication Type	Average Tutee Age	Reading Outcome Type	Assessment Name	Weeks	Sessions	Hours	Tutee Sample Size	Control Sample Size
1	Almaguer 2005(a)	Dyad	Same-Age	Journal	8.5	Fluency	Comprehensive Reading Assessment Battery, Fluency	9	45	22.5	40	40
2	Almaguer 2005(b)	Dyad	Same-Age	Journal	8.5	Comprehension	Comprehensive Reading Assessment Battery, Comprehension	9	45	22.5	40	40
3	Brown et al., 2018(a)	Dyad	Same-Age	Journal	8.5	Comprehension	Lexile	19	95	23.8	21	25
4	Brown et al., 2018(b)	Dyad	Same-Age	Journal	8.5	Comprehension	Lexile	19	95	23.8	18	25
5	Brown et al., 2018(c)	Dyad	Same-Age	Journal	8.5	Comprehension	Lexile	19	95	23.8	17	25
6	Brown et al., 2018(d)	Dyad	Same-Age	Journal	8.5	Fluency	DIBELS-ORF	19	95	23.8	21	25
7	Brown et al., 2018(e)	Dyad	Same-Age	Journal	8.5	Fluency	DIBELS-ORF	19	95	23.8	18	25
8	Brown et al., 2018(f)	Dyad	Same-Age	Journal	8.5	Fluency	DIBELS-ORF	19	95	23.8	17	25
9	Brown et al., 2018(g)	Dyad	Same-Age	Journal	8.5	Fluency	DIBELS-Acc	19	95	23.8	21	25
10	Brown et al., 2018(h)	Dyad	Same-Age	Journal	8.5	Fluency	DIBELS-Acc	19	95	23.8	18	25
11	Brown et al., 2018(i)	Dyad	Same-Age	Journal	8.5	Fluency	DIBELS-Acc	19	95	23.8	17	25
12	Burdett, 1986	PR	Adult	Journal	9.5	Comprehension	Wide Span	8	40	3.3	16	16
13	Cook et al., 1980(a)	NIM	Adult	Journal	10	Fluency	SRI Oral Reading	15	30	7.5	10	10
14	Cook et al., 1980(b)	NIM	Adult	Journal	10	Comprehension	SRI Silent Reading	15	30	7.5	10	10
15	Cook et al., 1980(c)	NIM	Adult	Journal	10	Fluency	SRI Oral Reading	6	30	5.0	12	12
16	Cook et al., 1980(d)	NIM	Adult	Journal	10	Comprehension	SRI Silent Reading	6	30	5.0	12	12
17	Diaper,1988(a)	PR	Cross-Age	TD	8.8	Comprehension	ERT, Comprehension of Sequences	8	40	10.0	15	15

(*table continues*)

Effect Number	Effect Name	Method	Tutor Type	Publication Type	Average Tutee Age	Reading Outcome Type	Assessment Name	Weeks	Sessions	Hours	Tutee Sample Size	Control Sample Size
18	Diaper,1988(b)	PR	Cross-Age	TD	8.8	Comprehension	ERT, Comprehension of Sequences	8	40	10.0	15	15
19	Diaper,1988(c)	PR	Same-Age	TD	8.8	Comprehension	ERT, Retention of Significant Ideas	8	40	10.0	15	15
20	Diaper,1988(d)	PR	Cross-Age	TD	8.8	Comprehension	ERT, Retention of Significant Ideas	8	40	10.0	15	15
21	Diaper,1988(e)	PR	Same-Age	TD	8.8	Comprehension	ERT, Use of Context	8	40	10.0	15	15
22	Diaper,1988(f)	PR	Cross-Age	TD	8.8	Comprehension	ERT, Use of Context	8	40	10.0	15	15
23	Diaper,1988(g)	PR	Same-Age	TD	8.8	Fluency	ERT, Rate	8	40	10.0	15	15
24	Diaper,1988(h)	PR	Cross-Age	TD	8.8	Fluency	ERT, Rate	8	40	10.0	15	15
25	Diaper,1988(i)	PR	Same-Age	TD	8.8	Comprehension	ERT. Comprehension of Essential Ideas	8	40	10.0	15	15
26	Diaper,1988(j)	PR	Cross-Age	TD	8.8	Fluency	ERT. Comprehension of Essential Ideas	8	40	10.0	15	15
27	Downs et al., 2020(a)	Dyad	Same-Age	Journal	8.5	Fluency	Acadience WCPM	12	60	15.0	82	20
28	Downs et al., 2020(b)	Dyad	Same-Age	Journal	8.5	Fluency	Acadience Accuracy	12	60	15.0	82	20
29	Downs et al., 2020(c)	Dyad	Same-Age	Journal	8.5	Comprehension	Acadience Retell	12	60	15.0	82	20
30	Downs et al., 2020(d)	Dyad	Same-Age	Journal	8.5	Comprehension	STAR	18	90	22.5	82	20
31	Eldredge& Quinn,1988	Dyad	Same-Age	Journal	7.5	Comprehension	Gates-MacGinite	36	180	NR	32	29
32	Joscelyn,1991(a)	PR	Same-Age	TD	10.25	Fluency	NEALE Accuracy	7	28	9.3	11	10
33	Joscelyn,1991(b)	PR	Same-Age	TD	10.25	Comprehension	NEALE Comprehension	7	28	9.3	11	10
34	Joscelyn,1991(c)	PR	Same-Age	TD	9.58	Fluency	NEALE Accuracy	7	28	9.3	11	11

(*table continues*)

Effect Number	Effect Name	Method	Tutor Type	Publication Type	Average Tutee Age	Reading Outcome Type	Assessment Name	Weeks	Sessions	Hours	Tutee Sample Size	Control Sample Size
35	Joscelyn,1991(d)	PR	Same-Age	TD	9.58	Comprehension	NEALE Comprehension	7	28	9.3	11	11
36	Joscelyn,1991(e)	PR	Same-Age	TD	10.25	Fluency	NEALE Accuracy	6	24	6.5	13	13
37	Joscelyn,1991(f)	PR	Same-Age	TD	10.25	Comprehension	NEALE Comprehension	6	24	6.5	13	13
38	Joscelyn,1991(g)	PR	Same-Age	TD	10	Fluency	NEALE Accuracy	4	20	6.7	11	11
39	Joscelyn,1991(h)	PR	Same-Age	TD	10	Comprehension	NEALE Comprehension	4	20	6.7	11	11
40	Joscelyn,1991(i)	PR	Same-Age	TD	9.75	Fluency	NEALE Accuracy	4	20	6.7	11	11
41	Joscelyn,1991(j)	PR	Same-Age	TD	9.75	Comprehension	NEALE Comprehension	4	20	6.7	11	11
42	Joscelyn,1991(k)	PR	Same-Age	TD	9.58	Fluency	NEALE Accuracy	6	30	7.5	11	11
43	Joscelyn,1991(l)	PR	Same-Age	TD	9.58	Comprehension	NEALE Comprehension	6	30	7.5	11	11
44	Joscelyn,1991(m)	PR	Same-Age	TD	9.75	Fluency	NEALE Accuracy	6	28	7.0	11	11
45	Joscelyn,1991(n)	PR	Same-Age	TD	9.75	Comprehension	NEALE Comprehension	6	28	7.0	11	11
46	Joscelyn,1991(o)	PR	Same-Age	TD	9.75	Fluency	NEALE Accuracy	6	28	7.0	11	11
47	Joscelyn,1991(p)	PR	Same-Age	TD	9.75	Comprehension	NEALE Comprehension	6	28	7.0	11	11

(*table continues*)

Effect Number	Effect Name	Method	Tutor Type	Publication Type	Average Tutee Age	Reading Outcome Type	Assessment Name	Weeks	Sessions	Hours	Tutee Sample Size	Control Sample Size
48	Joscelyn,1991(q)	PR	Same-Age	TD	10.08	Fluency	NEALE Accuracy	5	25	8.3	12	12
49	Joscelyn,1991(r)	PR	Same-Age	TD	9.75	Comprehension	NEALE Comprehension	5	25	8.3	12	12
50	Low et al., 1987	PR	Cross-Age	Conference	7.35	Comprehension	McLure Reading Test	4	20	5.0	13	13
51	Townsend, 1986(a)	PR	Same-Age	Conference	NA	Fluency	NEALE Accuracy	8	40	NR	6	9
52	Townsend, 1986(b)	PR	Same-Age	Conference	NA	Comprehension	NEALE Comprehension	8	40	NR	6	9
53	Weiss et al., 1989(a)	PR	Adult	TD	8.77	Comprehension	BASIS RAW	11	40	16.7	9	8
54	Weiss et al., 1989(b)	PR	Adult	TD	8.77	Fluency	CBM RAW	11	40	16.7	9	8
55	Young et al., 2015(a)	R2I	Adult	Journal	9.5	Fluency	DIBELS-ORF	20	20	6.7	29	23
56	Young et al., 2015(b)	R2I	Adult	Journal	9.5	Fluency	MDFS	20	20	6.7	29	23
57	Young et al., 2015(c)	R2I	Adult	Journal	9.5	Comprehension	iStation	20	20	6.7	29	23
58	Young et al., 2018a	R2I	Adult	Journal	7.98	Comprehension	iStation	6	18	6.0	25	25
59	Young et al., 2018b (a)	R2I	Adult	Journal	7.5	Comprehension	Passage Questions	7	21	7.0	19	18
60	Young et al., 2018b (b)	R2I	Adult	Journal	7.5	Fluency	MFS	7	21	7.0	19	18
61	Young et al., 2018b (c)	R2I	Adult	Journal	7.5	Fluency	Bader RLI	7	21	7.0	19	18

(*table continues*)

Effect Number	Effect Name	Method	Tutor Type	Publication Type	Average Tutee Age	Reading Outcome Type	Assessment Name	Weeks	Sessions	Hours	Tutee Sample Size	Control Sample Size
62	Young et al., 2018b (d)	R2I	Adult	Journal	7.5	Fluency	Bader RLI	7	21	7.0	19	18
63	Young et al., 2018b (e)	NIM	Adult	Journal	7.5	Comprehension	Bader RLI	7	21	7.0	20	18
64	Young et al., 2018b (f)	NIM	Adult	Journal	7.5	Fluency	MFS	7	21	7.0	20	18
65	Young et al., 2018b (g)	NIM	Adult	Journal	7.5	Fluency	Bader RLI	7	21	7.0	20	18
66	Young et al., 2018b (h)	NIM	Adult	Journal	7.5	Fluency	Bader RLI	7	21	7.0	20	18

Hedge's G Effect Size Calculation

The meta-analysis used to interpret these data measured two variations of growth from Paired Oral Reading: growth of the treatment group relative to growth of the control group and raw differences for treatment groups pre- and post- intervention. Accordingly, two variations of Hedge's g effect sizes were calculated for the selected data. First, an effect size was calculated for each of the 66 effects that utilized treatment/comparison designs. These effects measured the growth of treatment groups compared to growth of control groups. The treatment/comparison effect size data were used for all reported analyses except for a single sub-analysis in Research Question #1.

A second set of Hedge's g effect sizes was calculated for all 87 effects in the sample. These effects measured the raw growth of treatment samples, thus accounting for studies that measured gains, but did not include a control group. Importantly, these effects do not account for confounding variables such as maturation as well as the treatment/control effect sizes and must be interpreted accordingly. This second set of effect sizes was used for the sub-analysis mentioned in the previous paragraph. All other analyses utilized the treatment/control effect sizes. All effect sizes were calculated in R Studio using the escalc function from the metafor (Viechtbaur, 2010) package. Figure 4.1 displays the Hedge's g effect sizes for the 66 treatment/comparison effects. Figure 4.2 displays the Hedge's g effect sizes for the 33 of 66 treatment/comparison effects coded as measuring fluency outcomes. Figure 4.3 displays the Hedge's g effect sizes for the 33 of 66 treatment/comparison effects coded as measuring comprehension outcomes.

Figure 4.1

Hedge's g Effect Sizes for all Treatment/Control Outcomes

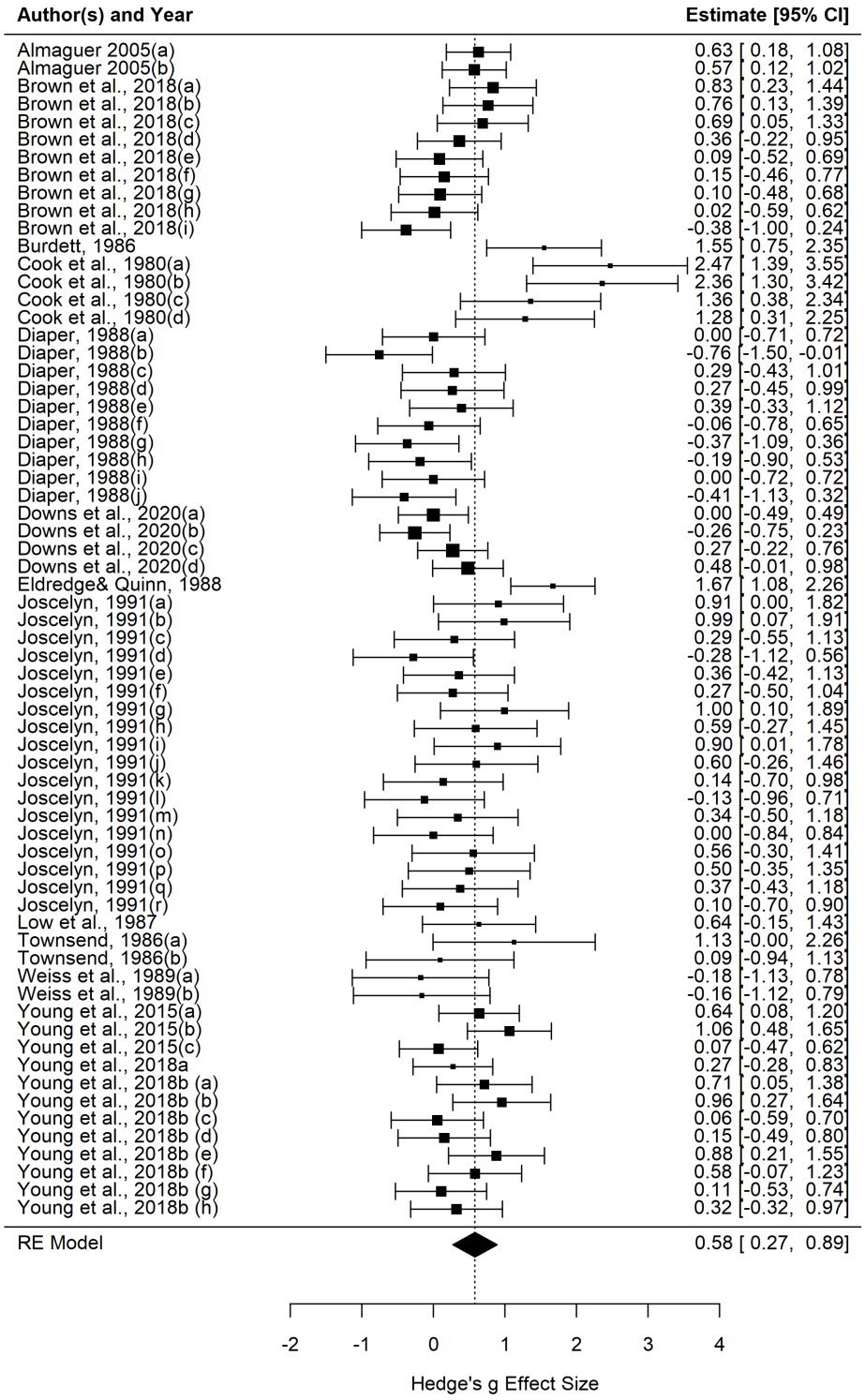


Figure 4.2

Outcomes for Treatment/Control Effects in Fluency

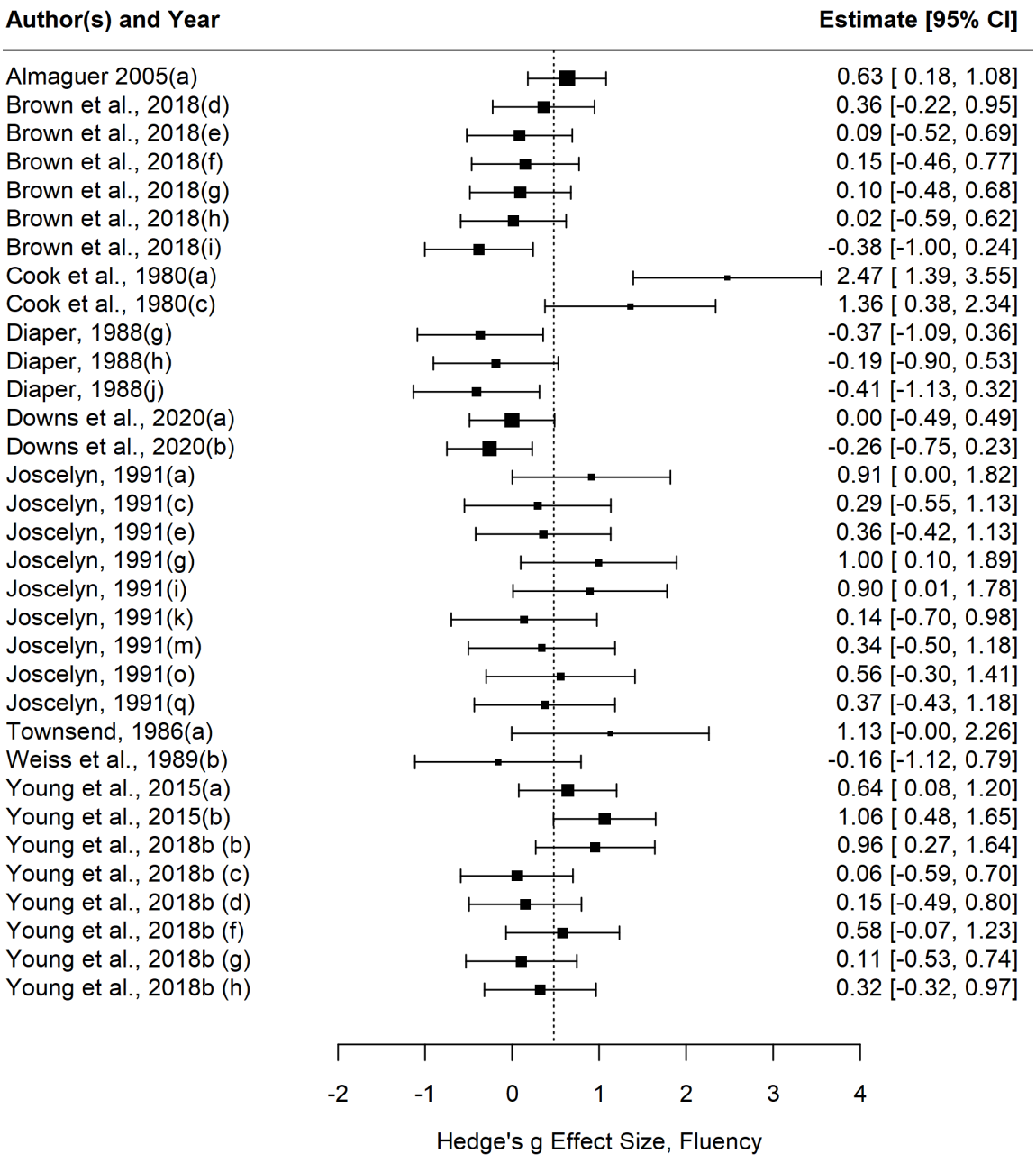
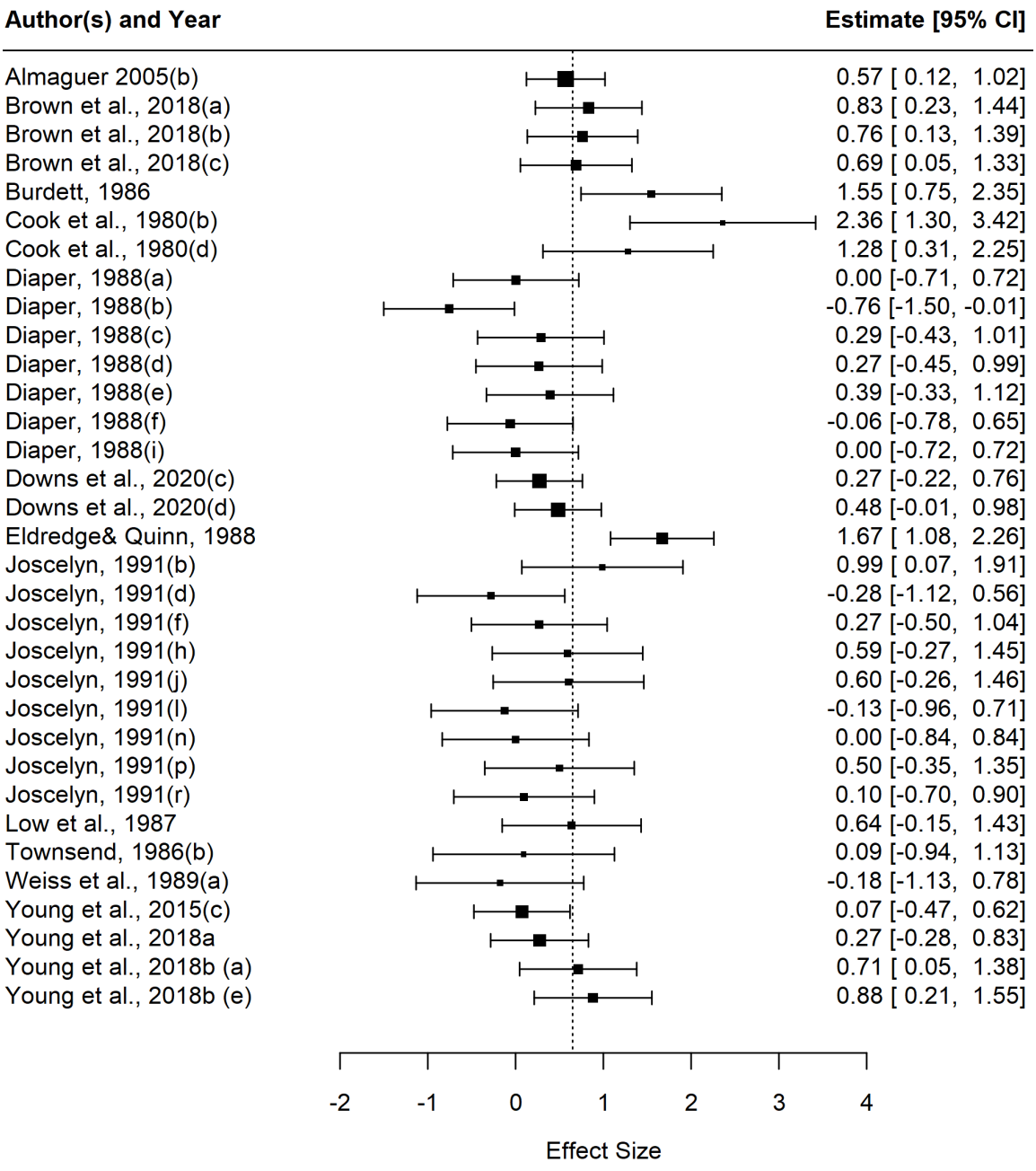


Figure 4.3

Outcomes for Treatment/Control Effects in Comprehension



Assessing Initial Model Fit

Between Study Heterogeneity

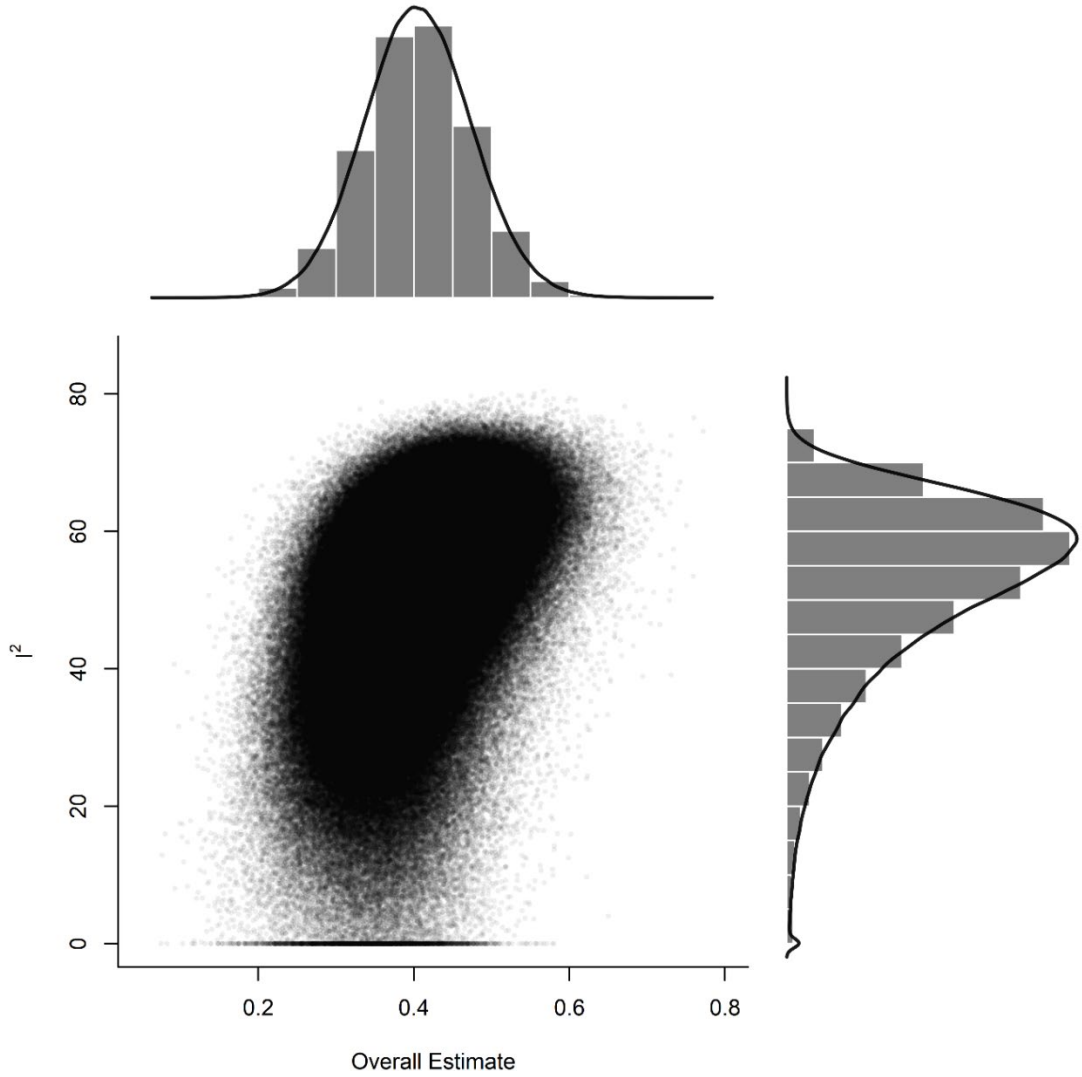
A three-level model was fit using the treatment/control effect sizes to determine between-study heterogeneity in the sample. This fit modeled effect sizes as level one, covariance among effect sizes within studies as level two, and covariance among studies as level three. This model reported a Q statistic of 143.56 ($p < .01$) and an I^2 percentage of 71.05%, indicating high heterogeneity between studies (Higgins et al., 2003; Jackson, 2013). The majority of heterogeneity was found at the study level (66%), followed by the effect level (28.95%). Only 5.05% of the I^2 was found among effects within studies. These results suggest that the studies included in the analysis varied considerably in design and outcomes. The high degree of heterogeneity present in the sample prompted using an innovative and robust method—graphical display of study heterogeneity (GOSH; Olkin et al., 2012)—to further investigate potential outliers or overly influential effects.

Figure 4.4 displays the results of the one million iterations of GOSH analysis. The plot displays several visual trends. First, the scatter plot appears to contain two clusters. The majority of effects appear to comprise a single large cluster in the middle of the graph. This main cluster that groups the majority of effects suggests that the influence among effects is relatively similar (Olkin et al., 2012). A second, minor cluster appears at the bottom of the graph in a thin line across the x-axis, indicating a range of effect size when the I^2 was at or near 0%. Also, the effect size histogram on the top of the plot appears evenly distributed, further suggesting homogeneity of influence among effects.

Last, the I^2 heterogeneity histogram on the right side appears to be positively skewed, indicating that heterogeneity of outcomes increased with effect size. Taken together, these data indicate high heterogeneity within the corpus of effect size, relatively similar influence on outcomes among the effect sizes, and a range of possible outcomes given the extant data.

Figure 4.4

Graphical Display of Study Heterogeneity



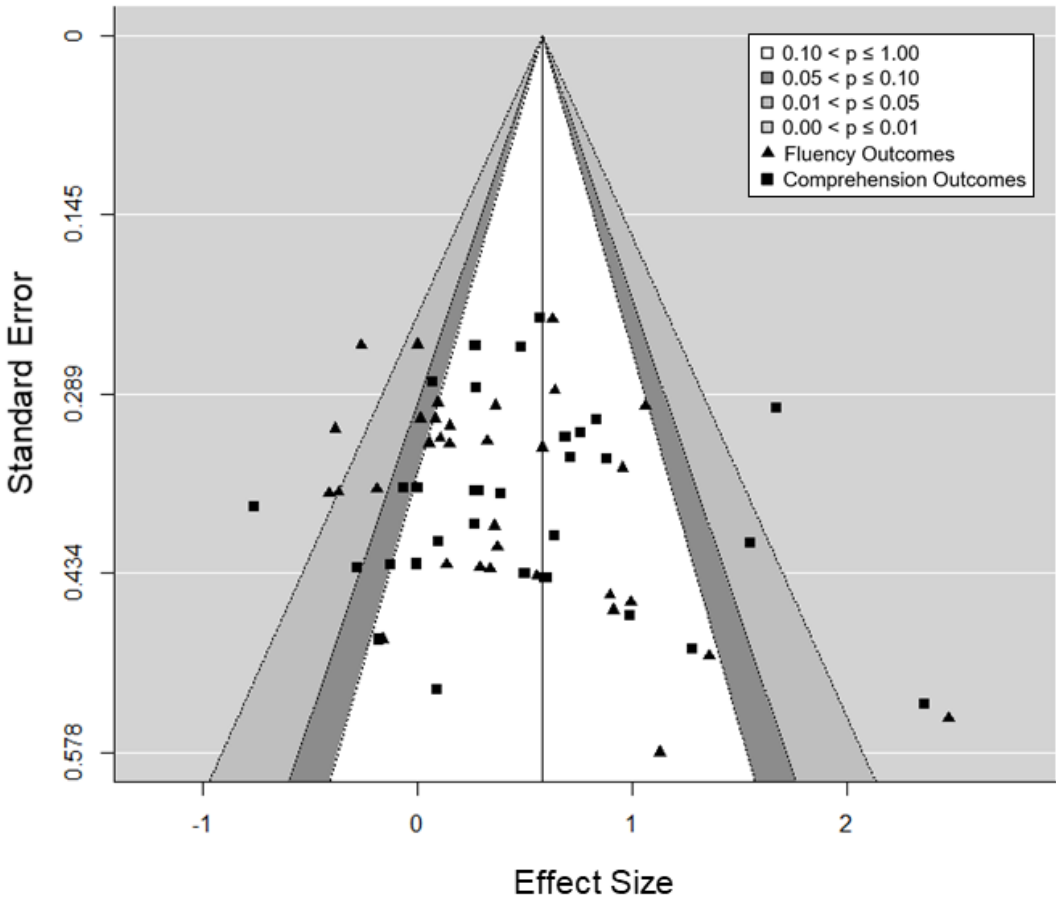
To further analyze the GOSH plot, three supervised machine learning algorithms (Gaussian Mixture Models, Fraley & Raftery, 2002; k-Means Clustering, Hartigan & Wong, 1979; DBSCAN, Schubert et al., 2017) were used to detect individual effects that contributed to cluster imbalance. These machine algorithms used density-based clustering and Cook's Distance (R. D. Cook, 1977) to identify effects that contributed to cluster imbalance. Each algorithm identified effects 1, 2, 3, 4, 6, 63, 64, 65, 66 (see Table 4.5) as contributing more to the GOSH outcomes relative to other outcomes in the corpus. Subsequently, these effects were removed from the analysis and the data were then refit into a three-level model. This reduced model reported a smaller degree of heterogeneity ($Q = 134.19$, $p < .01$) than the original three-level model ($Q = 143.56$, $p < .01$). However, this second model also reported increased variance ($\sigma = 0.485$) and I^2 statistic ($I^2 = 72.8\%$) than the original model ($\sigma = 0.424$; $I^2 = 71.05\%$. Although results from the machine learning algorithms suggest these effects appear to exert greater influence in the model relative to other effects, excluding them from the model does not appear to meaningfully improve study outcomes. Therefore, the original, more complex model was retained and investigated for publication bias.

Publication Bias

Publication bias was assessed using a funnel plot and modified Egger's Test (Hox et al., 2018; Lipsey & Wilson, 2001). Figure 4.5 displays a funnel plot of all treatment/control effects, coded as fluency and comprehension outcomes. The x-axis plots each observed effect size, and the y-axis plots the standard error for each effect. The displayed pattern appears mostly symmetrical, indicating the absence of publication bias.

Figure 4.5

Funnel Plot of Observed Outcomes



However, funnel plot interpretations can be problematic for multivariate analyses due to dependencies among multiple outcomes in a single study (Hox et al., 2018). Therefore, a modified Egger's test further investigated the potential for publication bias. Similar to a traditional Egger's test, the modified analysis can accommodate the multilevel, multivariate nature of the analysis. This test modeled the variance of each effect as a fixed effect within a full three-level model, as recommended by Hox et al. An omnibus test of the model coefficients reported statistically significant results (QM = 5.53, p =

0.02), suggesting the presence of publication bias in the data. Overall, these data suggest publication bias may be a limitation of the current literature in Paired Oral Reading.

Summary of Initial Fit

Several steps preceded answering the three research questions of this study. First, two sets of effect sizes were calculated given the coded data. These effect sizes compared the academic growth of treatment groups to that of the control groups followed by the aggregate growth of treatment groups. The treatment/control effect sizes were then fit to a three-level model. The model indicated a high degree of heterogeneity present in the data and a GOSH plot analysis was used to detect any potential outliers or overly influential effects. A reduced model that omitted detected outliers was then compared with the original model and the original model was found to be a better fit. Subsequently, publication bias was investigated using a funnel plot and modified Egger's test. While the funnel plot suggests the absence of publication bias, the modified Egger's test indicates that presence of publication bias. This model answers Research Question #1 and provides the comparison model for Research Question #2 and Research Question #3.

Results by Research Question

Research Question #1

Research question #1 asked, "*What is the effect of Paired Oral Reading on reading outcomes for tutees in grades K-6*"? The original three-level model enabled calculation of the effect of Paired Oral Reading on reading outcomes for tutees in grades K-6. This model used the treatment/comparison Hedge's *g* effect sizes as the first level,

the second level modeled the covariance between-effects and within studies, and the third level modeled the covariance across studies. This model was fit using restricted maximum likelihood (REML; Hox et al., 2018) with a 95% confidence interval. Results indicated a statistically significant effect of g = 0.58 (*SE* = .16, [95% CI = .27-.89]) for the 66 treatment/comparison effects.

Three more models were fit to render a more nuanced understanding of the relevant variables. The first two calculated the effect of Paired Oral Reading based on tutor type and academic outcome. For these, each variable was plotted on the original model as a fixed effect with the intercept removed using REML estimation. Outcomes for tutor type indicated statistically significant outcomes for adult (g = 0.73, *SE* = 0.25, [95% CI = 0.25–1.21], $p < 0.01$) and peer tutors (g = 0.52, *SE* = 0.21, [95% CI = 0.11 - 0.94], $p = .01$). Cross age tutors produced the smallest effect size (g = 0.26, *SE* = 0.28, [95% CI = -0.30 - 0.81]) and was not statistically significant ($p = .36$). Effect size by academic outcome indicated statistically significant results ($p < .01$) for fluency (g = 0.48, *SE* = 0.16, [95% CI = 0.16 - 0.80]) and comprehension (g = 0.65, *SE* = 0.16, [95% CI = 0.34–0.97], $p < .01$). The third model calculated the effect of all treatment group outcomes (n = 87) reported a large effect size of 1.32 (*SE* = .26, [95% CI = .82–1.82], $p < .01$). This model suggests the degree of academic growth that Paired Oral Reading yields when used as a classroom practice across a variety of conditions. The results indicate that outcomes are greater for adult tutors over peer or cross-age tutors and favors comprehension growth over fluency. Table 4.6 displays the results of the statistical analysis from Research Question #1.

Table 4.6

Results from Research Question #1

Model Fit	Effects	g	95% CI	SE	logLik	*p*	Q
All treatment effects	87	1.32	0.82, 1.82	0.26	-93.19	< .01	313.22
Treatment/control effects	66	0.58	0.27, 0.89	0.16	-44.08	< .01	143.56
Tutor type					-41.68		124.56
Adult	66	0.73	0.25, 1.21	0.25		< .01	
Cross-age	66	0.26	-0.30, 0.81	0.28		.36	
Peer	66	0.52	0.11-0.94	0.21		.01	
Reading outcome					-43.72		139.95
Fluency	66	0.48	0.16, 0.80	0.16		< .01	
Comprehension	66	0.65	0.34, 0.97	0.16		< .01	

Research Question #2

Research Question #2 asked, *"How do related factors (e.g., tutor type, dosage, year, publication type, and method used) moderate the reported outcomes"?* The heterogeneity identified during the initial assessment of model fit warranted further moderator analyses. Meta-regression techniques (Pigott & Polanin, 2020) were used to investigate any potential moderators. These moderator analyses sought to understand the degree that specific variables influence outcomes in the model. This method involved adding a single variable to the original model as a fixed factor. The moderator model was then compared to the original model using the likelihood ratio test (Hox et al., 2018). This process was repeated for each of the potential moderator variables. Results from the meta-regression moderator analysis are displayed in Tables 4.7–4.8.

Statistically Significant Moderator

Duration of intervention in weeks was the only moderator that produced a

Table 4.7

Results from the Research Question #2: Meta-Regression, Continuous Variables

Variable	*b*	*SE*	95% CI	*p*
Duration	0.04**	0.02	0, 0.07	.04
Dose	-0.01	0.02	-0.05, 0.03	.53
Sessions	0	0	0, 0.01	.16
Age	0.24	0.19	-0.13, 0.6	.21
Year	-0.02	0.01	-0.04, 0	.09

Table 4.8

Results from the Research Question #2 Meta-Regression, Categorical Variables

Variable	*b*	*SE*	95% CI	logLik	*p*
Reading outcome				3.02	.08
Fluency	0.48	0.16	0.16, 0.80		
Comprehension	0.65	0.10	0.34, 0.97		
Tutor type				2.08	.35
Adult tutor	0.73	0.25	0.25, 1.21		
Cross-age tutor	0.26	0.28	-0.30, 0.81		
Peer tutor	0.52	0.21	0.11, 0.94		
Tutoring method				1.62	.66
NIM	0.87	0.44	0.21, 1.52		
Dyad	0.61	0.29	0.04, 1.18		
Read Two Impress	0.72	0.42	0.11, 1.32		
Paired reading	0.42	0.39	-0.08, 0.93		
Publication type				3.81	.15
Journal	0.75	0.50	-0.39, 1.12		
Thesis/dissertation	0.07	.56	-0.54, 0.69		
Conference	0.60	0.46	-0.30, 1.51		
Study quality				2.20	.14
Random design	0.33	.30	-0.12, 0.78		
Non-random design	0.76	0.2	-0.12, 0.78		

statistically significant model (LRT = 4.21, $p < .04$). This model estimates that each one-week increase of the Paired Oral Reading intervention predicts the effect size increasing by .04 units ($b = .04$, $SE = .02$, $p = .04$). This outcome suggests that Paired Oral Reading regimens with a longer overall duration produce greater reading achievement outcomes for tutees.

Moderators Approaching Statistical Significance

Two models approached statistical significance: achievement outcome type (LRT = 3.02, $p = .08$) and publication year (LRT = 3.00, $p = .08$). The achievement outcome model estimated an effect size difference of 0.17 between fluency ($g = 0.48$, SE = 0.10, $p = .08$) and comprehension ($g = 0.65$, SE = 0.16, $p < .01$). Curiously, this model indicates a larger effect for Paired Oral Reading on comprehension outcomes than fluency ones. This model was nearly identical to the academic outcome model produced for Research Question #1; however, by retaining the intercept, this model analyzed whether the entire category explained more variance than the original three-level fit. This model indicated that although differences exist between fluency and comprehension outcomes, these differences do not moderate overall outcomes in Paired Oral Reading. The publication year model indicated that newer studies predict slightly lower outcomes; a one-year increase in publication is associated with a .02 decrease in effect size ($b = -0.02$, $SE = 0.01$, $p = .09$). This outcome may be tied to differences in sample sizes between older and more recent studies. Table 4.7 displays the meta-regression results of the continuous variables and Table 4.8 displays the meta-regressions results of the categorical variables used in the analysis.

Moderators Without Statistical Significance

Seven models with an added moderator variable produced no statistical significance: publication type (LRT = 3.81, p = .15), study quality (LRT = 2.70, p = .14), participant age (LRT = 1.58, p = .21), tutoring method (LRT =1.62, p = 0.66), total hours of intervention (LRT = 0.39, p = 0.53), and number of tutoring sessions (LRT = 1.95, p =.16). The intercept was also added to the tutor type model from Research Question #1 and was not statistically significant (LRT = 2.08, p = 0.35).

Research Question #3

Research Question #3 asked, *"What are the effects of Paired Oral Reading for tutees while accounting for significant moderator variables"?* This question was answered using two multiple meta-regression models derived from the original three-level model. The models followed a stepwise meta-regression format where multiple pre-determined covariates—tutor type and reading outcomes—were introduced into the model as fixed effects in a series. Therefore, these variables were included as part of the initial multiple meta-regression model. As outlined in Chapter 3, the second model retained the two predictors from the first model, while adding the only other statistically significant variable—duration—as a third fixed effect.

Multicollinearity Check

Multicollinearity was evaluated for the tutor type, academic outcome type, and duration of weeks variables using the variance inflation factor (VIF; James et al., 2013). The results indicated low collinearity among the covariates, allowing for their inclusion

in the multiple meta-regression.

Meta-Regression Model 1: Tutor Type and Achievement Outcome Type

The literature review in Chapter 2 identified tutor type and reading outcome as important variables in Paired Oral Reading, therefore these were included in both multiple meta-regression models. The first model was fit with the same nested structure as the original model from Research Question #1, adding tutor type and academic outcome type as fixed effect covariates. This multiple meta-regressed model explained more variance than the original model from Research Question #1 (LRT = 5.23, p = .16) using the Log Likelihood Ratio test; however, the outcome was not statistically significant.

Meta-Regression Model 2: Tutor Type, Achievement Outcome Type, and Duration in Weeks

The second multiple meta-regressed model added duration to the variables in the first model as a fixed effect. This model explained significantly more variance using the Log Likelihood Ratio test than the original fit informing Research Question #1 (LRT = 9.47, p = .05), and first meta-regressed model (LRT = 4.24, p = .04), suggesting this model as the most accurate fit of the analysis. Outcomes for this model are displayed in Table 4.9. This final model indicates that when controlling for duration, the estimated effect size of an adult tutor is 0.44 (*SE* = 0.30, [95% CI = -0.15 - 1.03], p = .14) for comprehension outcomes and 0. 28 (*SE* = 0.10, [95% CI = -0.30 – 0.86]), p = 0.35). for fluency outcomes. Further, each week of intervention added predicts an increase of 0.03

to the effect size. The results from same age tutors were less than adult tutors with an estimated effect of -0.02 (*SE* = 0.31, [95% CI = -0.15 – 1.03]) for fluency and 0.14 (*SE* = 0.30, [95% CI = -0.45 – 0.74]). The wide confidence intervals in these data suggest the analysis may be underpowered. Table 4.9 displays the coefficient results from the second multiple meta-regression model organized by fluency and comprehension outcomes.

Table 4.9

Coefficient Results from Multiple Meta-Regression Model #2

Coefficient	*b*	*SE*	*p*
Fluency outcomes			
Adult tutor	0.28	0.30	.35
Cross-age tutor	-0.23	0.36	.15
Same-age tutor	-0.02	0.31	.34
Duration	0.03	0.02	.05
Comprehension outcomes			
Adult tutor	0.44	0.30	.14
Cross-age tutor	-0.07	0.36	.15
Same-age tutor	0.14	0.30	.64
Duration	0.03	0.02	.05

Note. Coefficients are calculated when duration = 0 weeks. Coefficients are organized by fluency and comprehension outcomes.

Limitations

This dissertation represents a robust analysis of a common classroom practice using innovative methods. Multiple analyses and models were utilized to assess validity of the models and their outcomes. Notwithstanding, these findings are limited in several ways. First, these results are limited by the available data. Overall, the corpus of studies identified in the literature review represents an overall lack methodological rigor. It is

unknown to what degree the analysis is influenced by nonrandom study designs, although the moderator analysis did not indicate study quality as a statistically significant predictor. However, as explained next, the moderator analysis itself may have lacked power to detect differences in study quality.

Overall, there appears to be sufficient data to predict outcomes from Paired Oral Reading. However, the primarily null findings informing Research Question #2 combined with the large standard errors and confidence intervals throughout the analysis suggest an overall lack of statistical power. This potential limitation could be mitigated by future research that measures the fluency and comprehension outcomes for elementary-age tutees in Paired Oral Reading schemes using larger samples.

Omitted data constitute another limitation. Due to the screening protocol, the majority of effects represented complete data. However, some data were imputed by the researchers based on reasonable estimates, most commonly standard deviation information estimated using extant data (e.g., Joscelyne, 1991; A. Morgan et al., 2000; Topping & Whitely, 1993). Certain variables in some reports could not be reasonably imputed and were left blank in the data frame (e.g., Eldredge & Quinn 1988; Townsend, 1987).

Further, the findings reported here are potentially limited by the reading assessments used to measure tutee academic growth. The NARA has been critiqued frequently for various validity issues (Pumfrey, 1984; Stothard & Hulme, 1991; Topping 1990), yet this assessment constitutes the most common growth measurement used in the sample. The potential for publication bias (Card, 2015) is also present, given the

statistically significant result from the modified Egger's test. If publication bias is indeed present in the current sample, then outcomes from this analysis could potentially be inflated.

Critically, the corpus data may also limit the accuracy of the reported effect sizes. This meta-analysis compared the reading achievement growth of Paired Oral Reading tutees against the reading achievement growth of students in control groups. This type of comparison contains two types of dependencies: dependency between treatment and control group and a dependency between pre- and postscores. Accounting for both types of dependencies would be enhanced by correlation coefficients. However, correlation coefficients were rarely reported in the eligible data, and thus not collected in the coding process. As a result, effect size and effect size variance contain inaccuracies as the current model assumes $r = 0$, when in fact there is some degree of variance between pre- and post-assessment. While the lack of correlation coefficients may influence the accuracy of the calculated effect sizes, the overall pattern of results and relations between various models is likely unimpaired.

Other limitations include those potentially induced by the researcher. This dissertation, undertaken by a novice researcher, attempts to meld four separate lines of research across nearly six decades using complex, innovative methodologies. Moreover, the literature review, screening, coding, model-fitting, and interpretations of this study each contain intricate procedures that are nuanced and complex. Thus, the potential for human error or bias is present, albeit unintentional.

Summary of Findings

All studies meeting the inclusion criteria were dual coded as reported or imputed effects. The screening and coding process yielded 87 effects across 25 studies. The majority of these effects (n = 66) resulted from studies that utilized treatment/control type designs. A Hedge's g effect size was calculated for each independent outcome in the treatment/control cohort. Following heterogeneity and GOSH plot analyses, these data were used to fit a three-level model. The model reported an effect size of g = 0.58, suggesting a positive effect for Paired Oral Reading. Weeks of intervention duration was the only statistically significant moderator. A multiple meta-regression including tutor type, academic outcome type, and intervention duration explained significantly more variance than the original model, suggesting the best fit. These data indicate robust outcomes for tutees who participate in Paired Oral Reading schemes and suggest positive outcomes among a variety of implementations. Implications from this analysis—as delineated in the next chapter—include the viability of multilevel techniques in future literacy meta-analyses, recommendations for practitioners looking to support readers with Paired Oral Reading techniques and establishing an agenda for future research to elucidate the potential and challenges associated with implementation of Paired Oral Reading to support weaker elementary-grade readers.

CHAPTER 5

DISCUSSION AND CONCLUSION

Summary of Multilevel Outcomes

This meta-analysis reports positive results for tutees who engaged in Paired Oral Reading. Two main models were fit, one reporting the effect size of Paired Oral Reading tutees compared to students in a control group (g = .58), and one reporting the aggregate growth of tutees only between pre- and posttest measures (g = 1.32). Additional models comparing treatment and control outcomes indicated the relative superiority of adult tutors (g = .73), over same-age (g = .52) and cross-age (g = .26) peers and higher outcomes in reading comprehension (g = .65) than fluency (g = .48). Further analyses revealed that only duration of the intervention in weeks significantly moderated outcomes. Two other variables, reading achievement outcome type (LRT = 3.02, p = .08) and publication year (b = -.02, SE = .01, p = .09) approached statistical significance.

Lastly, two models were fit with multiple meta-regression to estimate the effect of Paired Oral Reading while accounting for significant variables. The first model, which measured outcome effects while controlling for tutor type and reading achievement outcome type, did not account for statistically significant more variance than the original model (LRT = 5.23, p = .16). This final multiple meta-regressed model calculated the effect of Paired Oral Reading while controlling for tutor type, reading outcome, and duration in weeks. The model accounted for significantly more variance than the original model (LRT = 9.47, p = .05) and the first meta-regressed model (LRT = 4.24, p = .04).

Similar to other models in the analysis, the second meta-regressed model predicted higher outcomes for comprehension than fluency and higher outcomes for adult tutors than peer tutors while controlling for duration in weeks of the intervention. The analyses isolated relatively few factors among many that contribute to the robust findings in the selected corpus of studies, which affords a focused discussion of instructional and research implications.

Discussion of Multilevel Outcomes

Synchronous Partnered Oral Reading Techniques

This dissertation investigated whether four methods of Paired Oral Reading–Paired Reading, Neurological Impress Method, Dyad Reading, and Read Two Impress–are similar enough to be considered within a single family of research. Results from the loglikelihood ratio test indicate that outcomes plotted by individual method did not produce a more accurate model, suggesting that that these four methods are indeed similar enough to be researched, analyzed, and reported together. This is an important finding from the various iterations of Paired Oral Reading that could inform future variations that may evolve. As such, a superordinate term is proposed to facilitate future research and application of these methods: Synchronous Partnered Oral Reading Techniques (SPORT). Consistent with practices scrutinized in this analysis, SPORT is defined as *a collection of tutoring techniques where a more proficient adult or peer regularly reads connected text orally and synchronously with a lower reading level student or peer with the goal to accelerate oral-reading development of the lower reader.*

Effects of Synchronous Partnered Oral Reading Techniques

Results from the current meta-analysis indicate robust outcomes for SPORT tutees. All of the statistically significant models reported effect sizes above Hattie's recommended $d = 0.40$ threshold for zone of desired effects (Fisher et al., 2016; Hattie 2009), and the majority of the effects are considered large using Cohen's (1988) estimation. The reported findings support the use of SPORT within classroom settings and as a subject for further empirical inquiry. Further, the sophisticated methodologies used in this analysis facilitate a nuanced interpretation of the effects of Paired Oral Reading from previous studies, which merits attention from educational researchers.

Effect of SPORT Among Investigated Variables

Fluency and Comprehension

An important goal of this dissertation was to ascertain the effect of Paired Oral Reading on fluency and comprehension outcomes. The model that fit fluency and comprehension as separate outcomes reported a higher effect size for comprehension-related outcomes ($g = 0.65$, $SE = 0.16$, $p < .01$) than for fluency outcomes ($g = 0.48$, $SE = 0.16$, $p < .01$). The moderator analysis of these outcomes approached statistical significance (LRT = 3.02, $p = .08$), suggesting these differences may warrant inspection. The I^2 statistic of the three-level model also provides evidence that outcomes may be higher for comprehension than fluency. Only 5.05% of the variance in the model occurred at the second level of the model, indicating that differences between outcomes were very consistent across studies. Essentially, the difference between fluency and

comprehension outcomes varied very minimally across the studies in the sample, suggesting that comprehension may indeed be influenced more by SPORT than fluency.

This finding is curious; SPORT provides oral reading practice for the tutee (Eldredge & Quinn, 1988; Embrey., 1968; A. Morgan et al., 2000), which would imply that fluency-related outcomes would be higher than comprehension outcomes. Indeed, educators typically implement SPORT to primarily support fluency development among pre-adolescent readers (Downs et al., 2020; Young et al., 2015). However, the evidence suggests that the opposite may be more likely; SPORT may support comprehension outcomes more than fluency outcomes. This counterintuitive finding could be explained in several different ways. First, fluency outcomes may be suppressed by ceiling effects in the data. Fluency is typically described as being comprised of rate, accuracy, and prosody (Kuhn & Stahl, 2003). This study did not code fluency outcomes as individual subcomponents, however post hoc review intimated that accuracy-related measures comprised the majority of the fluency sample. Oral reading accuracy has de facto ceiling of 100% and falls below 90% in only the poorest of readers (Hiebert, 2015) suggesting restricted potential of growth when measured simply as increases in word-reading accuracy. In other words, the narrow parameter for growth provided by measuring oral reading accuracy may help to explain the lower effect sizes for fluency. Future post hoc analyses of these and other data could help clarify different outcomes for rate, accuracy, and prosody measures to tease out any comparative differences within fluency and between fluency and comprehension.

A second explanation for the greater effect of Paired Oral Reading on

comprehension than fluency outcomes may lie in the practice itself. SPORT researchers (Brown et al, 2018; Downs et al., 2020; Topping & Lindsay, 1992a; Young, Pearce, et al., 2018). have frequently invoked LaBerge and Samuels' (1976) text processing framework to describe how these methods "free [readers] from the decoding burden and speed up the decoding process so that they can give the necessary attention to the text message" (Eldredge & Quinn, 1988, p. 45). The current meta-analytic results may support this theory; SPORT readers might be attending more to comprehension than fluency due to the tutored support. If this notion were valid, it would join a growing body of research suggesting that reading comprehension can be effectively promoted in developing readers by lessening or removing decoding demands (Baker et al., 2020; Mohr et al., 2021; Reutzel et al., 2016). Effectively, SPORT may indeed help readers "focus on prosody and meaning rather than decoding, speeding up word recognition and…syntactic parsing rather than decoding individual words" (Downs et al., 2020, p. 121). This possibility gives SPORT added benefit; not only does Paired Oral Reading appear to support the development of smooth, accurate readers, it may also help develop more proficient meaning makers.

Tutor Type

The meta-analytic findings suggest that adult tutors promote stronger academic outcomes than same-age tutors. Although these differences did not moderate overall outcomes, the notion is logical. An adult tutor likely possesses more competent reading skill and maturity than a tutee's peers, and thus may be more capable of providing productive tutoring sessions. The indicated differential effects of adult and peer tutors

come with an important caveat for research and practice; peer tutors are likely more widely available in school settings than adult tutors. Future research could investigate systems of SPORT implementation that wisely allocate peer and adult tutors to support lower achieving readers. Such a design could prioritize adult tutors for those students with greatest need and provide peer support for other students with less need. The number of effects utilizing cross-age tutors is likely too small to generalize any findings and should be further investigated with studies of larger sample size and increased rigor.

Method

As indicated in meta-regressed moderator analysis, the specific method of Paired Oral Reading did not predict different effect size outcomes. This ecological finding indicates that SPORT can be implemented in various ways to produce desired student outcomes. Although the method variable was not statistically significant overall, individual methods were statistically significant when plotted with the intercept removed. Considering differences among effect-size methods could help illustrate the affordances of each method. The Neurological Impress Method corresponded with the highest effect size (g = .87, SE = 0.44, [95% CI = 0.21–1.52]), followed by Read Two Impress (R2I; g = .72, SE = .42, [95% CI = 0.11–1.32]), Dyad Reading (g = .61, SE = .29, [95% CI = .04–1.18]), and Paired Reading (g = .42, SE = .39, [95% CI = -0.08–0.93]). These findings are consistent with the results from the tutor type model; NIM and R2I which use only adult tutors produced the highest effect sizes, while Dyad Reading, which only uses peer tutors, and Paired Reading, which uses a mixture of adult and peer tutors produced the smaller effect sizes. Importantly, the large standard errors and wide

confidence intervals suggest that when considered individually, each method was likely underpowered within the analysis. Further research in each of these methods would likely provide better estimates of the actual effects of these individual practices and their overall influence on SPORT outcomes.

Age of Tutee

Age of assisted reader did not significantly predict the outcomes of Paired Oral Reading. This finding is counterintuitive as acquiring reading competency is a developmental process, so it would thus make sense that students at certain ages would experience greater benefit from extended oral reading practice. Although exact reasons are unknown, the extant research occurred primarily within a narrow age range ($M = 8.9$, $SD = 0.9$), potentially precluding more nuanced findings of an optimal age for SPORT tutoring. This conclusion suggests that previous researchers already considered students within this age range primed for accelerated reading achievement via SPORT. Perhaps identifying the most appropriate population for SPORT is best done by considering students' aggregate reading achievement, rather than specific age. Students with decoding skill, but who lack word recognition, fluency, and comprehension may be best served by SPORT, however further research should test this notion.

Time Variables

Three variables measured the volume of paired reading that produced each effect: dose in hours, duration in weeks, and the total number of sessions. Despite the clear correlations among these variables, the duration of an intervention in weeks was the only

time moderator that proved statistically significant. This finding suggests that the overall length of an intervention may play the most important time-variable role in determining reading outcomes, suggesting a consistency of intervention may be more beneficial than the intensity of the intervention. Therefore, practitioners may choose to prioritize the number of weeks a student participates in SPORT over the number of sessions or minutes per session.

Results from the second meta-regressed model indicate that an adult-tutored intervention with a duration mean 10.2 weeks could produce an effect size of 0.75 in comprehension outcomes and 0.59 in fluency outcomes. This represents a substantial effect in under three months. The regressed model further predicts that each additional week of intervention would add .03 units to the effect size. These findings could support future researchers and practitioners in estimating the weeks needed to support desired growth for participating readers. However, it would also be advantageous to know if there is a point in duration where the potential effects of SPORT begin to diminish or include any deleterious effects on students (e.g., decreased interest in oral reading, impacted reader identities, or treatment fatigue).

Dosage of intervention in hours and total number of sessions did not significantly predict Paired Oral Reading outcomes. Further, not only were the models not statistically significant, the coefficients produced by the models approached zero (Dosage, b = -.01, SE = .02, [95% CI = -0.05–0.02]; Sessions, b = 0.0, SE = 0, [95% CI = 0.00–0.01]) The lack of finding for hourly dosage and total number of sessions is counterintuitive. If Paired Oral Reading has an effect, that effect must be related to the dosage of

intervention received, and thus the total number of sessions. These outcomes suggest lack of power in the model.

However, one interpretation of this finding is that, within the extant data, Paired Oral Reading was successfully implemented using varying hours of intervention and number of overall sessions. This suggests that other factors, including duration of the intervention in weeks, tutor availability, and tutee achievement level, should be considered when determining an appropriate dosage. Decisions on intervention dosage will likely be most effective when used within one standard deviation from the mean for hourly dosage (M = 11.5 hours, SD = 6.6) and number of sessions (M = 43.1, SD = 30.0) reported in the extant data. Taken together, these data suggest SPORT occurring for 15 minutes a session, 4 days a week, for 10 weeks will likely be sufficient for substantial growth in tutee reading outcomes. However, Paired Oral Reading appears to allow considerable flexibility in how it is implemented. Curiously, this dose of Paired Oral Reading is significantly lower than reported in the recent research in Dyad Reading (Brown et al., 2018; Downs et al., 2020) but similar to the dose reported in the recent R2I research (Young, Durham, et al., 2018; Young, Pearce, et al, 2018). Future research in Dyad Reading could incorporate lower, or varying, doses to investigate reading achievement outcomes.

Implications for Research and Instruction

This meta-analysis suggests several important implications beyond the reporting of effect sizes of Paired Oral Reading. The methodology utilized innovative multilevel meta-analytic techniques that are rapidly increasing but underused in contemporary

research (Fernández-Castilla et al., 2020). Moreover, the literature review and subsequent analyses synthesized previously disparate lines of research and attempts to meld them into a single field. This ambitious effort supports important implications for researcher and researcher and reading teacher alike.

Future Meta-analysis in Literacy Research

Multivariate Three-Level Models

This dissertation used innovative methodologies such as multilevel structures, multivariate analysis, graphic displays of heterogeneity, machine learning algorithms, and multiple meta-regressions to discern the effects of Paired Oral Reading. Traditional meta-analyses have used a two-level regression structure to calculate effect sizes. This analysis would have been severely crippled with this approach. Firstly, the multilevel analysis facilitated multiple effect sizes per study. The importance of this multifaceted approach cannot be understated. With a two-level structure and traditional analysis, only one effect size per study would be allowed, limiting the analysis to only 23 measured outcomes in the treatment/control analysis. The available literature on Paired Oral Reading is clearly multivariate with nearly every study reporting multiple reading achievement outcomes. Thus, a traditional meta-analysis would have allowed only sub-optimal options, such as averaging multiple outcomes within a study, selecting only one outcome per study, attempting to estimate the variance between outcomes, or running multiple univariate analyses (Hox et al., 2018). Each of these decisions would have necessitated discarding available data, severely limiting the meta-analytic outcomes.

This analysis added an intermediary level of regression that modeled the variance

between outcomes and allowed multiple effect sizes to be included from a single study. Thus, the three-level structure of the analysis approximately tripled the number of effects (n = 66 Treatment/Control; n = 87 Treatment Only) that could be included in the analysis. The three-levels of regression used in this meta-analysis comprised a critical component of the research design and greatly magnified what is now known about the collective research in Paired Oral Reading.

Similarly, the three-level model allowed for a multivariate analysis, which facilitated the finding that Paired Oral Reading may produce greater outcomes for comprehension measures than fluency. With a two-level-single-outcome-per-study approach, fluency and comprehension outcomes could not have been analyzed concomitantly because the majority of studies contained at least one fluency and one comprehension outcome. Again, this limitation would provide only sub-optimal choices. Foremost, estimating different outcomes for fluency and comprehension outcomes might have been omitted altogether. Another option could have been to use two separate two-level analyses; one for fluency and one for comprehension. These two separate meta-analyses would then be compared to estimate any differences. However, comparing outcomes between these two separate analyses would have been limited due to the inherent covariance between the modeled effects, and, relatedly, the data discarded to construct each model (Hox et al., 2018). The added intermediary level allowed for fluency and comprehension outcomes to be modeled in a single fit that accounted for their covariance, greatly enhancing the accuracy of how these two outcomes are related and providing a nuanced finding of how SPORT influences reading achievement.

One last major methodological implication of the three-level model involves the I^2 heterogeneity statistic. This statistic indicated large heterogeneity at the study-level ($I^2_{Level\ 3} = 66\%$) and the effect-level ($I^2_{Level\ 1} = 28.95\%$). but near homogeneity at the between-effect level ($I^2_{Level\ 2} = 5.05\%$). These data suggest that while variables such as method, tutor type, dose, and the assessment used greatly influenced the outcomes, outcomes within single studies were remarkably consistent. Consistency between effects suggests concomitant consistency of SPORT outcomes, despite very heterogenic circumstances. Put simply, even though the outcomes between studies varied, outcomes within studies were consistent, suggesting the efficacy of SPORT tutoring regimens. The nuanced view of the between-effect heterogeneity would be lost in a two-level model, limiting the outcomes and interpretations of this analysis.

Meta-Regression and Multiple Meta-Regression

This analysis also used meta-regression and multiple meta-regression to determine the effects of Paired Oral Reading. The moderator analysis in this research design used meta-regressive techniques to determine the influence of specific variables on SPORT outcomes. Typical approaches to meta-regression add a covariate as a single fixed factor to the model and then assess statistical significance using an omnibus test of factors. The result is a list of variables that moderate outcomes and a list of variables that do not. This analysis used a model-building approach to determine significant moderator variables as recommended by Hox et al. (2018). In this approach, the new model with the added covariate was compared to the original model using a log likelihood test to determine which model best fit the data. While the omnibus test is appropriate for a two-

level analysis, the log likelihood approach is more appropriate for the three-level structure, contributing to the validity of the meta-regressed outcomes (Hox et al., 2018). Further, fitting a new model for each moderating variable allowed an effect size to be estimated for that variable, whether or not the outcome was statistically significant. This allowed for a more refined interpretation of the outcomes that would not have been possible with a more traditional significant/not significant approach to moderator analysis.

Multiple meta-regression is not a new technique in meta-analysis; however, evidence suggests that this technique is underused in contemporary research (Tipton et al., 2019a) The major benefit of this technique is that multiple meta-regression controls for multiple variables, reporting a more accurate outcome of the overall effect size. A meta-analysis is essentially a null model that pools effect sizes into a single outcome, accounting for the nested structure. Multiple meta-regression can add covariates to the model similar to other regression techniques, which predict and control for those variables. This is important because variables of interest within meta-analysis are usually related and estimating results of variables in relation to one another is typically desirable. Essentially, meta-analysis reports a pooled effect size, meta-regression reports an effect size while controlling for a single variable, and multiple meta-regression reports an effect size while controlling for multiple variables. Thus, multiple meta-regression is likely to capture more of the variance in the model than either meta-analysis or meta-regression, thereby optimizing the use of data within the model.

Future Meta-Analytic Designs in Literacy Research

Incorporating the innovative techniques noted above proved fruitful in the current analysis. Future meta-analytic designs in literacy should work to incorporate methods such as multilevel models, multivariate analysis, multiple meta-regressed effect sizes, and other robust techniques. These advanced methodologies align well with the types of data literacy researchers are interested in collecting. Literacy research often measures multiple outcomes from study participants along with various student and environmental characteristics. Multilevel methodologies can provide more nuanced multivariate analyses, and multiple meta-regression can help account for variables of interest within the meta-analytic model. Constructing research designs that optimize the available data will likely accelerate what is known about literacy practices and outcomes.

Implications for Classroom Practice

Close reading of the relevant research and the subsequent meta-analyses inform the use of SPORT to support reader development and afford specific practices to enhance efficacy. What follows relies on the particular details of SPORT as understood by the previous literature review and outcomes from the meta-analysis. Further, these outcomes are proposed within a Response to Intervention (RtI; Gersten et al., 2008) framework to support contextual fit within individual school sites.

Designing SPORT Interventions

Tutee participation in SPORT schemes should be contingent on student need, based on pre-determined grade-level benchmarks in oral reading fluency. Screening

procedures using curriculum-based measures common to RtI practices should suffice to identify students who may benefit from SPORT tutoring. These students would likely be in Grades 2 and 3, however upper-elementary students who remain dysfluent may benefit from these practices. Text complexity was not explicitly included as a variable in the meta-analysis; however, the literature review revealed several studies that report accelerated outcomes for tutees reading texts above their independent reading level (Downs et al., 2020; Flood et al., 2005; Young, Pearce, et al. 2018). Specifically, two grade levels above independent reading level of the tutee appears provide optimal benefit (Brown et al., 2018; A. Morgan et al., 2000). Tutees should be given a range of books to select from with this selection continually adapted to maintain appropriate text difficult as the student gains proficiency (Brown et al., 2018). Importantly, research-based recommendations on text complexity on SPORT outcomes are currently tentative, but flexible, so student progress should be monitored through the duration of the intervention.

Although comprehension outcomes appear to be greater than fluency outcomes, SPORT tutoring should not be utilized as a comprehension intervention or displace comprehension instruction. SPORT primarily consists of oral reading support and practice for developing readers. This oral reading development may serve as a vehicle for increased meaning making, but any comprehension growth is indirect. Summarily, SPORT is not comprehension instruction.

Some previous iterations of SPORT have incorporated comprehension support within the tutoring regimen (Eldredge & Quinn, 1988; Flood et al., 2005). Tutoring schemes like this have potential for future investigation and use by practitioners, but the

potential for volunteer adult or peer tutors to provide robust comprehension instruction may be limited. Ultimately, the increased comprehension outcomes gained from SPORT are important and should be considered by researchers and practitioners alike, but the primary implementation of SPORT should target oral reading fluency development.

Length of Intervention

Despite the inconsistent findings related to volume and intensity, practitioners must consider the dose and duration when applying SPORT interventions. Results from the meta-analysis indicate that duration of the intervention predicted successful student outcomes and that various doses corresponded with successful outcomes. Teachers who implement SPORT using time variables one standard deviation from the mean are likely to yield successful student outcomes. This equates to approximately 15 minutes per session across 4-16 weeks, for a total 5-18 hours of tutoring, although interventions outside of this range have also been impactful (c.f., Almaguer 2005; Brown et al., 2018).

As with any instructional practice, dosage variables should be contingent on student need. Based on pre-determined academic benchmarks in the area of oral reading fluency students with higher need could receive a greater overall dose of intervention. This conception allows a more flexible implementation of Paired Oral Reading than the whole-class, one-size-fits-all, approaches reported in the research (e.g., Brown et al., 2018; Downs et al., 2020; A. Morgan et al., 2000). Flexible implementations may be better adapted to the dynamic realities of classroom instruction. For example, students could be screened for inclusion in SPORT schemes, and then progress-monitored during the intervention. As students achieve pre-determined benchmarks, SPORT support could

be faded, and then discontinued as proficiency stabilizes. Student progress should be monitored throughout this process so that students not responding to this intervention could receive more targeted, supplemental instruction, with or without SPORT. As outlined, screening for inclusion in, and progress monitoring during, SPORT interventions align closely with RtI frameworks (Gersten et al., 2008) and could support student achievement outcomes more strategically.

Using Tutors Effectively

Relevant results indicate that adults and higher achieving peers can be trained to use SPORT with weaker readers. Tutor selection could consider that adult tutors will likely have a greater effect than higher level peers. Given the probable limited availability of adult tutors compared to peer tutors, adult tutors could be reserved for those students with greater need, while those with lesser need are assigned peer tutors. The extant research reports adult tutors as typically comprising recruited community volunteers or trained school staff (Flood et al., 2005; Young Pearce, et al., 2018). These populations may be accessible for those wishing to implement versions of SPORT with adult tutors. Peer tutors are also likely available for strategic support of weaker readers. Close monitoring of peer-tutoring sessions could help partners optimize the available learning time. Further, some research suggests that peer tutors can experience burnout when tutoring for extended periods (Downs et al., 2020); therefore, peer tutors should be rotated or relieved at regular intervals, to minimize their time tutoring and, presumably, to allow them other productive academic opportunities.

Designing Class and School SPORT Interventions

Classroom. Within an individual classroom, SPORT could be implemented similarly to Eldredge and Quinn's (1988) conception. This scheme integrated dyad reading during regular classroom instruction as needed. Teachers selected students who would benefit from receiving support from a buddy reader. Lead readers then volunteered to provide support daily whenever independent reading was required. Although a presumably simple integration, a potentially large volume of tutoring could occur over time. This example highlights the flexibility of SPORT and shows the opportunity for contextual fit within individual classrooms.

Alternatively, a teacher may wish to implement SPORT as a short-term Tier II type intervention that supplements Tier I classroom instruction. In this RtI variation, SPORT could provide targeted practice for oral reading achievement. Regular progress monitoring could then measure reader progress, allowing for data-based decisions by the teacher and other stakeholders. If used as a formal Tier II intervention, a classroom teacher would likely want to utilize adult tutors to help ensure fidelity of the tutoring and to facilitate the greatest reading achievement growth.

School site. A single school site may wish to coordinate efforts for a single program, similar to those reported by Young, Durham, et al. (2018) and Young, Pearce, et al. (2018). This method could result in grade levels compiling a list of prospective tutees, and community outreach by the school to recruit volunteer adult tutors. Tutees could then be prioritized based on recorded achievement levels, with the highest-need tutees receiving SPORT support with adult tutors, while the remaining tutees receive peer

tutors. These students might then be progress-monitored and shifted from adult to peer tutor or released from tutoring altogether based on pre-determined levels of achievement, with the adult tutors always tutoring the students with greatest need. Such a scheme would perhaps take considerable coordination but could provide valuable supplemental instruction for those needing to make accelerated progress to attain grade-level benchmarks.

Summary of Implications for Practice

Ostensibly, Paired Oral Reading in various iterations has been successful under a variety of circumstances. Given the findings of this research, certain guidelines will likely enhance tutee SPORT outcomes. SPORT tutoring occurs with a higher-level reader regularly reading aloud synchronously with a lower-level reader with text sufficiently complex to challenge the tutee. Progress of tutees should be monitored and the tutoring adjusted accordingly to maintain ongoing progress in fluency, and indirectly, comprehension. Within these relatively simple requirements, myriad implementations of SPORT are possible, and indeed tenable, for practitioners and students. Despite the findings of the present study that extend previous research recommendations, more studies are needed to confirm and elucidate the power and potential of SPORT.

Agenda for Future SPORT Research

Conceptualizing SPORT tutoring as a family of related methods will likely greatly enhance the literature base and facilitate a more widespread knowledge about these techniques. The current literature review and subsequent meta-analyses suggest a

robust future agenda for SPORT research. The following sections delineate an ambitious agenda for SPORT researchers. This ambitious research trajectory encompasses a broad range of topics, including theories of change, synthesis across techniques, and translating the literature into practice.

Theories of Change

Despite nearly six decades of research, little is known about the actual mechanism of how Paired Oral Reading improves reading achievement. For example, over 30 years ago, methods Eldredge and Quinn (1988) speculated that dyad reading can "free [readers] from the decoding burden and speed up the decoding process so that they can give the necessary attention to the text message" (p. 45). However, these authors also note that "this theory is entirely speculative and must be tested by further research" (Eldredge & Quinn, 1988, p.45). Researchers over the years have embraced the decoding burden hypothesis (Brown et al., 2018; Downs et al., 2020; Flood et al., 2005; Topping & Lindsay 1992a; Young, Pearce, et al., 2018), and its concomitant link to LaBerge and Samuels (1986) theory of automaticity but have neglected investigations into the reasons SPORT influences reading outcomes. Essentially, the decoding burden hypothesis is as speculative today as it was in 1988. Future research could investigate the cognitive functions of Paired Oral Reading and establish a theory of change that supports its mechanism. These findings would likely enhance the empirical and practical implications of Paired Oral Reading.

If indeed SPORT lessens any decoding burden, such changes could potentially be investigated using eye-tracking technology. Eye tracking affords spatial (via scan

paths and heat maps) and temporal measures, such as pupil fixations, saccades, and regressions as proxies indicative of cognitive processing (Just & Carpenter, 1980). While tracking words across a page, pupils fixate longer on words that are more challenging compared to those that are simpler or more familiar. Consequently, fluctuations in pupil movement and fixations between independent and Paired Oral Reading contexts could substantiate the purported text processing benefits of Paired Oral Reading.

Future research supporting the decoding burden hypothesis could also explain the higher effect sizes in reading comprehension outcomes compared to fluency. Students who regularly struggle through text due to inefficient word reading may have fewer cognitive resources to devote to making meaning (LaBerge & Samuels, 1986). Repeated sessions of Paired Oral Reading may promote the tutee's text processing capability and accelerate the meaning making processes. Essentially, a lighter decoding load may support students to construct coherent mental models of text. If, due to underdeveloped decoding and fluency achievement, constructing coherent mental representations is inhibited, then progression of comprehension achievement could proceed at an unhindered rate through SPORT. This notion appears to be intimated in the data but needs further investigation. Such a conception of SPORT may be framed better within theories of reading comprehension such as Kintsch's (1988) Construction-Integration Theory or Perfetti's Lexical Quality hypothesis (Perfetti, 2007) than the more typical automaticity theory of LaBerge and Samuels. Although implications of the decoding burden hypothesis are intriguing, the decades-old statement by Eldredge and

Quinn (1988) applies; these theories and their implications are highly speculative, and merit further investigation.

The role of increased volume of reading could also be investigated concomitantly with any investigations of text processing benefits. Volume of word reading is an important predictor of reading achievement outcomes (Hiebert, 2015) and a potentially major benefit of SPORT tutoring. It may be that a primary influence of SPORT is simply that the tutee successfully reads complex, connected text at a greater volume than they could independently. Such a notion could frame SPORT within neo-Vygotskian conceptions of zone of proximal development (c.f., Young et al., 2015). Future research could investigate the role of increased volume of reading on student achievement outcomes in SPORT, while determining effective ways to control for its influence.

Any investigation into theory of change influenced by Paired Oral Reading should work to unify the theoretical frameworks these methods invoke. Specifically, R. Morgan's (1976) Paired Reading method was built upon and continues to be practiced using behaviorist principles of praise and corrections (c.f. Topping & Lindsay, 1992a; Tymms et al 2011). Although Paired Reading researchers acknowledge the involvement of a text processing mechanism of Paired Oral Reading (Topping & Lindsay, 1992a), the practice itself remains behavioristic. A clearer understanding of how SPORT support student reading outcomes could shift these behavioral responses to practices that align more completely with a more robust theory of change.

Synthesizing SPORT Research

This meta-analysis serves as a systematic review and synthesis of related Paired Oral Reading techniques and provides effect sizes as benchmarks for future research. Subsequent investigation on individual SPORT techniques would benefit from interpreting results within the greater SPORT literature. For example, a study using Paired Reading with peer tutors could evaluate outcomes with peer-tutor effect sizes from this analysis, or with results from the Dyad Reading literature. In addition to synthesizing the reported outcomes of SPORT, the techniques themselves can be informed from other practices. For example, Read Two Impress (Young et al., 2015) incorporates repeated reading as part of its tutoring regimen. A method like dyad reading could incorporate repeated reading in a similar manner, resulting in a 'Dyad Two Impress' technique. Such syntheses could lead to a modular approach to Paired Oral Reading; different characteristics of Paired Oral Reading being "stacked" (Mohr et al., 2012) contingent on student need. A modular conception of SPORT could further borrow other effective fluency and comprehension practices to enhance student outcomes.

Bridging SPORT into Effective Practice

Facilitating access to the SPORT literature for practitioners is perhaps the most pressing issue for contemporary SPORT researchers. Paired Oral Reading is unequivocally pragmatic; volunteer adult and peer tutors are available at most school sites and the tutoring regimen is relatively easy to train and implement. Further, this practice is extremely cost efficient, requiring only the range of books typically available at the school or local libraries and more proficient tutors. The combination of high effect

and low cost makes these practices attractive, and indeed productive, for many participants. Communicating this important research to practitioners may help them use these methods effectively within their classrooms.

One way to communicate SPORT tutoring to practitioners could be through diverse methodologies, such as formative and design experiments (Reinking & Bradley, 2008). Such research would begin with a pedagogical goal and then use cycles of data to measure progress. With practitioners and researchers as partners, such research would focus less on isolating individual variables for analysis and more on how to effectively implement SPORT within school sites. This approach to researching SPORT could minimize potential misinterpretations of the SPORT literature by well-intended practitioners and maximize the potential for contextual fit within an individual school site. Suggestions for future design research on SPORT include resurrecting the integrated approach explained by Eldredge and Quinn (1988) or exploring the school-wide model reported by Young, Pearce, et al. (2018). These experiments could utilize available adult and peer tutors contingent on student need and monitor their progress toward grade-level benchmarks. Such investigations could also utilize the modular approach proposed in the previous section.

After conducting this research, communicating outcomes and effective practices could help practitioners implement SPORT judiciously. Myriad opportunities exist for this communication, including practitioner journal articles, conference presentations, on-site professional development, podcast interviews, and other outlets. Ultimately, future research in Paired Oral Reading must do better to translate previous outcomes into

effective practice. The reported meta-analysis and future research endeavors will do little good if they are not communicated effectively and implemented more extensively by classroom practitioners.

SPORT Tutees

Future research should also investigate the influence of Paired Oral Reading on diverse populations, including students of varying levels of socioeconomic status, English language learners, students with documented Individualized Education Plans (IEPs), and disengaged readers. The rates and characteristics of these populations are poorly reported in the extant research and the outcomes of SPORT among these populations would represent a significant achievement within the literature. Further, populations of students should be investigated in the upper-grades and beyond, especially those with a history of inadequate fluency. Limited studies currently exist in these areas.

Dose and Duration

Optimal dose and duration of SPORT interventions have yet to be clarified. Future research must determine the differential outcomes of SPORT tutees based on overall dose and duration. The current model assumes a linear relationship between these time variables and effect size outcomes. However, the actual relationship may be curvilinear; an initial slow, gradual rise, followed by a sharp rise in achievement before leveling off over time. This theory may be supported by duration being the only time-related significant variable. Recent research has reported surges of growth based on text complexity and time variables (Brown et al., 2018; Downs et al., 2020). Further, S-curves

of growth such as the one suggested here have appeared elsewhere in fluency research (Sabatini et al., 2019). Future research should investigate the presence of any curvilinear relationships between time and outcomes while identifying estimates of optimal intervention duration and dose.

Peer Tutors

Recent research has reported no added benefits for tutors of SPORT schemes (Downs et al., 2020), countering claims by other studies and proponents (Topping 1990; Topping & Bryce, 2004; Topping & Lindsay 1992a). This outcome appears logical in that higher level peer readers primarily scaffold access to the text for the lower-level reader, while reading texts that may not challenge the more proficient partner. Thus, although peer tutors facilitate benefit for the assisted reader, they may not themselves benefit from the experience. Downs et al. recommended rotating peer tutors for SPORT tutees in an effort to minimize their dosage of tutoring. Future research may wish to further investigate this notion, perhaps by qualitatively capturing the experience of lead readers. Future research may also investigate characteristics of peer tutors that predict successful tutee outcomes.

Text Complexity

Recent conceptions of SPORT techniques have consistently utilized complex texts to facilitate tutor achievement (Brown et al., 2018; Downs et al., 2020; Flood et al., 2005 A. Morgan et al., 2000; Young, Pearce, et al., 2018). Given the current focus on complex text driven by the Common Core State Standards (National Governors

Association [NGO], 2010) and the success of tutees using texts above their independent level, future research should continue to incorporate complex text strategically. Although results appear to support the use of complex text, future research could investigate this notion, including monitoring presence of surges in achievement based on text complexity as reported in recent studies (Brown et al., 2018; Downs et al 2020).

Previous research has typically identified text complexity using text-level measures such as SRI Lexile or the Accelerated Reader STAR systems. Although these leveling measures are widely available at school sites and represent the pragmatic nature of SPORT, future research may wish to adopt a more nuanced estimation of text complexity. These studies could investigate measures of text cohesion on student achievement outcomes using sophisticated tools such as Coh-Metrix (Graesser et al., 2011). By measuring outcomes based on indices of text cohesion—such as those present in narrative versus informational text structures—researchers may be able to specify text-level predictors of accelerated achievement, and perhaps infer the mechanism of change provided by SPORT.

Reading Attitudes

This meta-analysis did not investigate the influence of SPORT on reading attitudes. However, the literature review revealed a long history of reading attitudes measured concomitant with SPORT outcomes. Some may value Paired Oral Reading because of its use of authentic, connected text and the opportunity for one-on-one tutored support. These characteristics imply that one outcome of SPORT could be improved reading attitudes. However, as indicated in the literature review, the relationship between

reading attitudes and SPORT remains unclear (Downs et al., 2020). Future research could investigate the degree of influence that Paired Oral Reading has on reading attitudes. Such investigations would benefit from including contemporary motivational or engagement theories to inform the analysis. Further, these investigations could analyze whether shifting reading attitudes are moderated by the degree of reading achievement growth. It may be that increases in oral reading achievement is a better predictor of improved reading attitude than participating in SPORT itself (McKenna et al., 1995). Lastly, the scales commonly used to assess reading attitude may need to be further investigated to ascertain their degree of sensitivity to reading attitude changes and to what degree these scales reflect the literacy lives of contemporary, 21st century readers.

Conclusion

A single premise initiated this dissertation study: should the varied forms of Paired Oral Reading be considered as a family of related practices? Synthesis from the literature review and meta-regression from the subsequent analysis appear to indicate that, indeed, these methods are similar enough in practice and outcome to be considered together. Forging these methods into a field—Synchronous Partnered Oral Reading Techniques— represents a significant shift in the literature of these four practices. This dissertation sought to report what is known about SPORT, report potential practices for school-site implementation, and establish a robust agenda for future SPORT research. Further, due to the innovative nature of the analysis, this dissertation challenges literacy meta-analytic researchers to produce more sophisticated and nuanced designs to better ascertain what is known about literacy outcomes. Enhanced meta-analytic designs,

thoughtful, pragmatic implementations of SPORT tutoring, and future research into these practices will hopefully provide incremental but important steps to supporting developing readers across grade levels and contexts.

REFERENCES

Alexander, P. A. (2020). Methodological guidance paper: The art and science of quality systematic reviews. *Review of Educational Research*, *90*(1), 6–23. https://doi.org/10.3102/0034654319854352

Allington, R. (1983). Fluency: The Neglected Reading Goal. *The Reading Teacher, 36*(6), 556-561. http://www.jstor.org/stable/20198272Almaguer, I. (2005). Effects of dyad reading instruction on the reading achievement of Hispanic third-grade English language learners. *Bilingual Research Journal*, *29*(3), 509–526. https://doi.org/10.1080/15235882.2005.10162849

Baker, D. L., Santoro, L., Biancarosa, G., Baker, S. K., Fien, H., & Otterstedt, J. (2020). Effects of a read aloud intervention on first grade student vocabulary, listening comprehension, and language proficiency. *Reading and Writing*, *33*, 2697-2724. https://doi.org/10.1007/s11145-020-10060-2

Barrett, J. (1986). Reading with Mother A paired reading programme described and evaluated. *Literacy*, *20*(3), 173–178. https://doi.org/10.1111/j.1467-9345.1986.tb00341.x

Barrett, J. M. (1987). Paired reading: Psycholinguistics in practice. *Literacy*, *21*(3), 152–158. https://doi.org/10.1111/j.1467-9345.1987.tb00638.x

Barrett, T. S., & Brignone, E. (2017). Furniture for quantitative scientists. *The R Journal*, *9*(2), 142-148. https://doi.org/10.32614/RJ-2017-037

Bedsworth, B. (1991). The neurological impress method with middle school poor readers. *Journal of Reading*, *34*(7), 564–565. https://www.jstor.org/stable/40014585

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97–111. https://doi.org/10.1002/jrsm.12

Borenstein, M., Higgins, J. P. T., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. *Research Synthesis Methods*, *8*(1), 5–18. https://doi.org/10.1002/jrsm.1230

Brown, L. T., Mohr, K. A. J., Wilcox, B. R., & Barrett, T. S. (2018). The effects of dyad reading and text difficulty on third-graders' reading achievement. *Journal of Educational Research*, *111*(5), 541–553. https://doi.org/10.1080/00220671.2017.1310711

Buck, J., & Torgesen, J. (2018). The relationship between performance on a measure of oral reading fluency and performance on the Florida Comprehensive Assessment Test (FCRR Technical Report# 1). *Florida Center for Reading Research*. https://eric.ed.gov/?id=ED587688

Burge, B., Styles, B., Brzyska, B., Cooper, L., Shamsan, Y., Saltini, F., & Twist, L. (2010). *New Group Reading Test (NGRT)*. GL Assessment.

Bus, A. G., Van Ijzendoorn, M. H., & Mol, S. E. (2011). Meta-analysis. In N. K. Duke & M. H. Mallette (Eds.), *Literacy research methodologies* (2nd ed., pp. 270–300). Guilford Press.

Bush, A. (1983). Can pupils' reading be improved by involving their parents. *Remedial Education*, *18*(4), 167–170.

Bushell, R., Miller, A., & Robson, D. (1982). Parents as remedial teachers. *Journal of the Association of Educational Psychologists*, *5*(9), 7–13. https://psycnet.apa.org/record/1983-11310-001

Cadieux, A., & Boudreault, P. (2005). The effects of a parent-child paired reading program on reading abilities, phonological awareness and self-concept of at-risk pupils. *Reading Improvement*, *42*(4), 224–237.

Card, N. A. (2015). *Applied meta-analysis for social science research* (1st ed.). Guilford Press.

Cawood, S., & Lee, A. (1985). Paired reading at Colne Valley High School. *Paired Reading Bulletin*, *1*, 46–50. https://files.eric.ed.gov/fulltext/ED285124.pdf

Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. Psychological Methods, 19(2), 211–229. https://doi.org/10.1037/a0032968

Cheung, M. W. L. (2019). A Guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychology Review* 29**,** 387–396. https://doi.org/10.1007/s11065-019-09415-6

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Erlbaum.

Cohen, P. A., Kulik, J. A., & Kulik, C.-L. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal*, *19*(2), 237–248. https://doi.org/10.3102/00028312019002237

Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics, 19*(1), 15-18. https://doi.org/10.1080/00401706.1977.10489493

Cook, J. E., Nolan, G. A., & Zanotti, R. J. (1980). Treating auditory perception problems: The NIM helps. *Academic Therapy*, *15*(4), 473–481. https://doi.org/10.1177/105345128001500410

Crombie, R., & Low, A. (1986). Using a paired reading technique in cross age peer tutoring. *Paired Reading Bulletin*, *2*(1), 10–15. https://files.eric.ed.gov/fulltext/ED285126.pdf

Daniel Lüdecke. (2018). *sjstats: Statistical functions for regression models.* Zenodo. https://doi.org/10.5281/zenodo.1489175

DeAngelo, N., Reents, K., & Zomboracz, C. (1997). *Improving reading achievement through the use of parental involvement and paired reading* [Master's thesis, Saint Xavier University]. https://eric.ed.gov/?id=ED409536

Deeks, J. J., Higgins, J. P., Altman, D. G. (2021). Analysing data and undertaking meta-analyses. *Cochrane handbook for systematic reviews of interventions*, 241-284. https://training.cochrane.org/handbook/current/chapter-10

Del Re, A. C. (2015). A practical tutorial on conducting meta-analysis in R. *The Quantitative Methods for Psychology*, *11*(1), 37–50. https://doi.org/10.20982/tqmp.11.1.p037

Diaper, G. R. (1989). *A comparative study of paired reading techniques using parent, peer cross-age tutors with second year junior school children* [Doctoral thesis, University of Kent at Canterbury]. https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.330228

Dixon-Krauss, L., Januszka, CynthiaM., & Chae, C.-H. (2010). Development of the dialogic reading inventory of parent-child book reading. *Journal of Research in Childhood Education*, *24*(3), 266–277. https://doi.org/10.1080/02568543.2010.487412

Downs, J. D., Mohr, K. A. J., & Barrett, T. S. (2020). Determining the academic and affective outcomes of dyad reading among third graders. *The Journal of Educational Research*, *113*(2), 120–132. https://doi.org/10.1080/00220671.2020.1755615

Dufrene, B. A., Reisener, C. D., Olmi, D. J., Zoder-Martell, K., McNutt, M. R., & Horn, D. R. (2010). Peer Tutoring for Reading Fluency as a Feasible and Effective Alternative in Response to Intervention Systems. *Journal of Behavioral Education*, *19*(3), 239–256. https://doi.org/10.1007/s10864-010-9111-8

Elbaum, B., Vaughn, S., Tejero Hughes, M., & Watson Moody, S. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology*, *92*(4), 605–619. https://doi.org/10.1037/0022-0663.92.4.605

Eldredge, J. L. (1988a). A 52-year-old dyslexic learns to read. *Journal of Reading, Writing, and Learning Disabilities International*, *4*(2), 101–106. https://doi.org/10.1080/0748763880040204

Eldredge, J. L. (1988b). Sacred cows make good hamburger. *Academic Therapy*, *23*(4), 375–382. https://doi.org/10.1177/105345128802300408

Eldredge, J. L. (1988c). Improving the reading comprehension skills of poor readers. *Reading Horizons: A Journal of Literacy and Language Arts*, *29*(1), 35-42. https://scholarworks.wmich.edu/reading_horizons/vol29/iss1/4

Eldredge, J. L. (1990). Increasing the performance of poor readers in the third grade with a group-assisted strategy. *Journal of Educational Research*, *84*(2), 69–77. https://doi.org/10.1080/00220671.1990.10885995

Eldredge, J. L., & Butterfield, D. (1986). Alternatives to traditional reading instruction. *The Reading Teacher*, *40*(1), 32-37. http://www.jstor.org/stable/20199303

Eldredge, J. L., & Quinn, D. W. (1988). Increasing reading performance of low-achieving second graders with dyad reading groups. *Journal of Educational Research*, *82*(1), 40–46. https://doi.org/10.1080/00220671.1988.10885863

Ellis, M. G. (1995). *The effects of a parent-child reading program on reading ability and self-perceptions of reading ability in struggling young readers* [Doctoral dissertation, University of Maryland, College Park]. http://search.proquest.com/docview/304214918/abstract/2A6D4B461CFB4039PQ/1

Embrey, A. C. (1968). *A study on the effectiveness of neurological impress method as remedial reading technique* [Master's thesis, Central Washington State College]. https://digitalcommons.cwu.edu/etd/822

Esteves, K. J., & Whitten, E. (2011). Assisted reading with digital audiobooks for students with reading disabilities. *Reading Horizons: A Journal of Literacy and Language Arts*, *51*(1), 21-40. https://scholarworks.wmich.edu/reading_horizons/vol51/iss1/4

Ewing, J., & Johnstone, M. (1981). *Attitudes to reading: Measurement and classification within a curricular structure* [Unpublished Manuscript, Dundee College of Education].

Fernández-Castilla, B., Jamshidi, L., Declercq, L., Beretvas, S. N., Onghena, P., & Van den Noortgate, W. (2020). The application of meta-analytic (multi-level) models with multiple random effects: A systematic review. *Behavior Research Methods*. https://doi.org/10.3758/s13428-020-01373-9

Fiala, C. L., & Sheridan, S. M. (2003). Parent involvement and reading: Using curriculum-based measurement to assess the effects of paired reading. *Psychology in the Schools*, *40*(6), 613–626. https://doi.org/10.1002/pits.10128

Fisher, D., Frey, N., & Hattie, J. (2016). *Visible learning for literacy, grades K-12: Implementing the practices that work best to accelerate student learning*. Corwin Press.

Flood, J., Lapp, D., & Fisher, D. (2005). Neurological impress method PLUS. *Reading Psychology*, *26*(2), 147–160. https://doi.org/10.1080/02702710590930500

Fraley, C., & Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American statistical Association*, *97*(458), 611-631. https://doi.org/10.1198/016214502760047131

Fry, S. (2014). *An exploration of paired reading with a peer and its impact on the reading ability and school connectedness of looked-after children* [Doctoral thesis, University of Nottingham]. http://eprints.nottingham.ac.uk/27712/1/Stephanie%20Fry%20Thesis.pdf

Fuchs, D., Fuchs, L. S., & Burish, P. (2000). Peer-assisted learning strategies: An evidence-based practice to promote reading achievement. *Learning Disabilities Research & Practice*, *15*(2), 85–91.

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. https://doi.org/10.1207/S1532799XSSR0503_3

Gallagher, T. M. (2008). *The effects of a modified duet reading strategy on oral reading fluency* [Ph.D., The University of Wisconsin - Madison]. ProQuest Dissertations and Theses Global.

Gately, R. (2014). *An evaluation of a paired reading intervention implemented by foster carers with looked after children* [Doctoral thesis, University of Birmingham]. https://etheses.bham.ac.uk/id/eprint/5123/

Gautrey, F. (1988). Paired reading in a middle school. *Reading*, *22*(3), 175–179. https://doi.org/10.1111/j.1467-9345.1988.tb00674.x

Gersten, R., Compton, D. L., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2008). *Assisting students struggling with reading: Response to Intervention and multi-tier intervention for reading in the primary grades. A practice guide* (NCEE 2009-4045). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/wwc/publications/practiceguides/

Good, R. H., Kaminski, R. A., Smith, S., &. Laimon, D. (2002). *Dynamic indicators of basic early literacy skills: DIBELS*. Dynamic Measurement Group, Inc.

Good, R. H., Kaminski, R. A., Dewey, E. N., Wallin, J., Powell-Smith, K. A., & Latimer, R. (2011). *Dynamic Measurement Group, Inc*.

Goudey, J. (2009). *A parent involvement intervention with elementary school students: The effectiveness of parent tutoring on reading achievement* [Doctoral dissertation, University of Alberta]. https://era.library.ualberta.ca/items/2941ab77-bab7-480c-9ce3-e2340d068db3

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational researcher*, *40*(5), 223-234. https://doi.org/10.3102/2F0013189X11413260

Harrer, M., Cuijpers, P., Furukawa, T. & Ebert, D. D. (2019). *dmetar: Companion R package for the guide "Doing meta-analysis in R"* (R package version 0.0.9000). http://dmetar.protectlab.org/

Hartigan, J. A., & Wong, M. A. (1979). AK-means clustering algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *28*(1), 100-108. https://doi.org/10.2307/2346830

Hasbrouck, J., & Tindal, G. (2017). *An update to compiled ORF norms* [Technical Report No. 1702]. Behavioral Research and Teaching, University of Oregon. https://www.brtprojects.org/wp-content/uploads/2017/10/TechRpt_1702ORFNorms_Fini.pdf

Hattie, J (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Heathington, B. S. (1975). *The development of scales to measure attitudes toward reading* [Unpublished Manuscript]. University of Tennessee Knoxville.

Heckelman, R. G. (1966). Using the neurological impress remedial reading technique. *Academic Therapy Quarterly*, *1*(4), 235–250. https://doi.org/10.1177/105345126600100410

Heckelman, R. G. (1969). A neurological-impress method of remedial reading instruction. *Academic Therapy*, *4*(4), 277–282. https://doi.org/10.1177/105345126900400406

Heckelman, R. G. (1986). N.I.M. revisited. *Academic Therapy*, *21*(4), 411–420. https://doi.org/10.1177/105345128602100403

Hedges, L., & Olkin, I. (1985). *Statistical Methods for Meta-analysis* (1st ed.). Academic Press. https://doi.org/10.1016/C2009-0-03396-0

Henk, W. A. (1983). Adapting the NIM: To improve comprehension. *Academic Therapy*, *19*(1), 97–101. https://doi.org/10.1177/105345128301900118

Hermsmeyer, K. (1999). *Training low -income parents of first-grade students in paired reading: The effects on reading fluency and attitudes toward reading and school* [Doctoral dissertation, University of the Pacific]. ProQuest Dissertations and Theses Global.

Hewison, J., & Tizard, J. (1980). Parental involvement and reading attainment. *British Journal of Educational Psychology*, *50*(3), 209–215. https://doi.org/10.1111/j.2044-8279.1980.tb00803.x

Hiebert E. H. (2015) The forgotten reading proficiency: stamina in silent reading. In Hiebert E. H. (Ed.), *Teaching stamina and silent reading in the digital-global age* (pp. 16-31). Text Project.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ : British Medical Journal*, *327*(7414), 557–560. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC192859/

Hollingsworth, P. M. (1970). *An experiment with the impress method of teaching reading, 24*(2), 112–114. https://www.jstor.org/stable/20196454

Howell, A. (2008). Parent-assisted reading using a paired reading model. *The Corinthian*, *9*, 49–64. https://kb.gcsu.edu/thecorinthian/vol9/iss1/11

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2018). *Multilevel analysis: Techniques and applications* (3rd ed.). Routledge.

Jackson, D. (2013). Confidence intervals for the between-study variance in random effects meta-analysis using generalised Cochran heterogeneity statistics. *Research Synthesis Methods*, *4*(3), 220–229. https://doi.org/10.1002/jrsm.1081

James, G., Daniela, W., Trevor, H., & Robert, T. (2013). *An introduction to statistical learning: with applications in R*. Spinger.

Joscelyne, T. (1991). *Peer-tutored paired reading* [Doctoral dissertation, University of Sussex]. https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.332012

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological review*, *87*(4), 329. https://psycnet.apa.org/doi/10.1037/0033-295X.87.4.329

Kann, R. (1983). The method of repeated readings: Expanding the neurological impress method for use with disabled readers. *Journal of Learning Disabilities*, *16*(2), 90–92. https://doi.org/10.1177/002221948301600205

King, G. T. (1986). *The effects of a modified neurological impress method and paired reading on high school remedial readers' oral reading fluency, comprehension, and use of context* [Doctoral dissertation, University of Georgia]. ProQuest Theses and Dissertations Global.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. Psychological Review, 95(2), 163-182. https://doi.org/10.1037/0033-295X.95.2.163

Kintsch, W. (2018). Revisiting the construction–integration model of text comprehension and its implications for instruction. In D.E. Alvermann, N.J. Unrau, M. Sailors, & R.B. Ruddell (Eds.), *Theoretical Models and Processes of Literacy* (8th ed., pp. 194-219). New York: Routledge.

Klvacek, M. L., Monroe, E. E., Wilcox, B., Hall-Kenyon, K. M., & Morrison, T. G. (2017). Follow the reader: An effective strategy to support students reading more complex text. *Reading Psychology*, *38*(5), 542–551. https://doi.org/10.1080/02702711.2017.1310159

Klvacek, M. L., Monroe, E. E., Wilcox, B., Hall-Kenyon, K. M., & Morrison, T. G. (2019). How second-grade English learners experienced dyad reading with fiction and nonfiction texts. *Early Childhood Education Journal*, *47*(2), 227–237. https://doi.org/10.1007/s10643-018-0919-5

Kuhn, M. R., Schwanenflugel, P. J., & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly*, *45*(2), 230–251. https://doi.org/10.1598/RRQ.45.2.4

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, *95*(1), 3–21. https://doi.org/10.1037/0022-0663.95.1.3

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*(2), 293–323. https://doi.org/10.1016/0010-0285(74)90015-2

Lancy, D. F., & Nattiv, A. (1992). Parents as volunteers: Storybook readers/listeners. *Childhood Education*, *68*(4), 208–212. https://doi.org/10.1080/00094056.1992.10520876

Law, M., & Kratochwill, T. R. (1993). Paired reading: An evaluation of a parent tutorial program. *School Psychology International*, *14*(2), 119–147. https://doi.org/10.1177/0143034393142003

Leach, D. J., & Siddall, S. W. (1990). Parental involvement in the teaching of reading: A comparison of hearing reading, paired reading, pause, prompt, praise, and direct instructional methods. *British Journal of Educational Psychology*, *60*(3), 349–355. https://doi.org/10.1111/j.2044-8279.1990.tb00951.x

Lees, E. (1985). An account of a paired reading project. *The Paired Reading Bulletin*, *1*(1), 68–69. https://files.eric.ed.gov/fulltext/ED285124.pdf

Leung, K. C. (2019). An updated meta-analysis on the effect of peer tutoring on tutors' achievement. *School Psychology International*, *40*(2), 200–214. https://doi.org/10.1177/0143034318808832

Limbrick, E., McNaughton, S., & Glynn, T. (1985). Reading gains for underachieving tutors and tutees in a cross-age tutoring programme. *Journal of Child Psychology and Psychiatry*, *26*(6), 939–953. https://doi.org/10.1111/j.1469-7610.1985.tb00608.x

Lindsay, G., Evans, A., & Jones, B. (1985). Paired reading versus relaxed reading: A comparison. *British Journal of Educational Psychology*, *55*(3), 304–309. https://doi.org/10.1111/j.2044-8279.1985.tb02636.x

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE.

Lloyd, C., Edovald, T., Kiss, Z., Morris, S., Skipp, A., & Ahmed, H. (2015). *Paired reading: Evaluation report and executive summary*. Education Endowment Foundation. https://eric.ed.gov/?id=ED581127

Lorenz, L., & Vockell, E. (1979). Using the neurological impress method with learning disabled readers. *Journal of Learning Disabilities*, *12*(6), 420–422. https://doi.org/10.1177/002221947901200614

Low, A., & Davies, M. (1988). Infant/junior cross-age peer group tuition. *Paired Reading Bulletin*, *4*(1), 7–10.

Low, A., Madden, L., & Davies, M. (1987). Infant/junior cross-age peer group tuition. *Paired Reading Bulletin*, *3*(1), 91–95. https://files.eric.ed.gov/fulltext/ED285126.pdf

Mathes, P. G., Torgesen, J. K., Clancy-Menchetti, J., Santi, K., Nicholas, K., Robinson, C., & Grek, M. (2003). A comparison of teacher-directed versus peer-assisted instruction to struggling first grade readers. *Elementary School Journal*, *103*(5), 459–479. https://doi.org/10.1086/499735

Maxwell, S. E. (2004). The persistence of underpowered studies in psychological research: Causes, consequences, and methods. *Psychological Methods*, *9*(2), 147–163. https://doi.org/10.1037/1082-989X.9.2.147

McAllister, E. A. (1989). A study of peer tutors using the neurological impress method. *Proceedings of Annual Meeting of the Eastern Educational Research Association*, 1-15.

McKay, M. (1996). The Neale analysis of reading ability revised—Systematically biased? *British Journal of Educational Psychology*, *66*(2), 259–266. https://doi.org/10.1111/j.2044-8279.1996.tb01194.x

McKenna, M. C., Stratton, B. D., Grindler, M. C., & Jenkins, S. J. (1995). Differential effects of whole language and traditional instruction on reading attitudes. *Journal of Reading Behavior, 27*(1), 19–44. doi:10.1080/10862969509547867

McKenna, M. C., & Kear, D. J. (1990). Measuring attitude toward reading: A new tool for teachers. *The Reading Teacher*, *43*(9), 626–639. https://www.jstor.org/stable/20204872

Medcalf, J. (1989). Comparison of peer tutored remedial reading using the pause, prompt, and praise procedure with an individualised tape-assisted reading programme. *Educational Psychology*, *9*(3), 252–262. https://doi.org/10.1080/0144341890090307

Mellard, D. F., Anthony, J. L., & Woods, K. L. (2012). Understanding oral reading fluency among adults with low literacy: Dominance analysis of contributing component skills. *Reading and Writing*, *25*(6), 1345–1364. https://doi.org/10.1007/s11145-011-9322-y

Mikkelsen, V. P. (1981, November 13). *The effects of a modified neurological impress method on developing decoding skills*. East Carolina University Library Science Lecture Series, Greenville, NC. https://eric.ed.gov/?id=ED209638

Miller, A. (1985). The role of paired reading in home-school collaboration schemes. *Curriculum*, *7*(1), 33–38.

Miller, A. (1987). Is There Still a Place for Paired Reading? *Educational Psychology in Practice*, *3*(1), 38–43. https://doi.org/10.1080/0266736870030108

Miller, A. L., & Narrett, C. M. (1995). Does parent involvement and parent feedback about reading progress influence students' reading progress? *Proceedings of Annual Meeting of the American Psychological Association (103rd)*, 1-33.

Miller, A., Robson, D., & Bushell, R. (1986). Parental participation in paired reading: A controlled study. *Educational Psychology*, *6*(3), 277–284. https://doi.org/10.1080/0144341860060305

Miller, B. V., & Kratochwill, T. R. (1996). An evaluation of the paired reading program using competency-based training. *School Psychology International*, *17*(3), 269–291. https://doi.org/10.1177/0143034396173003

Mohr, K. A. J., Dixon, K., & Young, C. (2012). Effective and efficient: Maximizing literacy assessment and instruction. In E. Ortlieb & E. H. Cheek (Ed.s) *Using informative assessments towards effective literacy instruction* (Literacy Research, Practice and Evaluation, Vol. 1, pp. 293-324). Emerald Group Publishing Limited. https://doi.org/10.1108/S2048-0458(2012)0000001013

Mohr, K. A. J., Chou, P., Tsai, Z., & Downs., J. D. (2021). *Using read-and-think-alouds to teach inferencing: Supporting comprehension from the beginning* [Manuscript submitted for publication]. Department of Teacher Education and Leadership, Utah State University.

Morgan, A., Wilcox, B. R., & Eldredge, J. L. (2000). Effect of difficulty levels on second-grade delayed readers using dyad reading. *The Journal of Educational Research*, *94*(2), 113–119. https://doi.org/10.1080/00220670009598749

Morgan, R. (1976). "Paired reading" tuition: A preliminary report on a technique for cases of reading deficit. *Child: Care, Health and Development*, *2*(1), 13–28. https://doi.org/10.1111/j.1365-2214.1976.tb00230.x

Morgan, R., & Gavin, P. (1988). Paired reading: Evaluation and progress. *Support for Learning*, *3*(4), 201–206. https://doi.org/10.1111/j.1467-9604.1988.tb00098.x

Morgan, R., & Lyon, E. (1979). "Paired reading:" A preliminary report on a technique for cases of reading deficit. *Journal of Child Psychology and Psychiatry*, *20*(2), 151–160. https://doi.org/10.1111/j.1469-7610.1979.tb00495.x

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common core state standards for English language arts and literacy in history/ social studies, science, and technical subjects*. http://www.corestandards.org/wp-content/uploads/ELA_Standards1.pdf

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development. https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf

Neale, M. D. (1958). *Neale analysis of reading ability*. MacMillian.

Nickow, A., Oreopoulos, P., & Quan, V. (2020). *The impressive effects of tutoring on preK-12 learning: A systematic review and meta-analysis of the experimental evidence*. National Bureau of Economic Research. https://doi.org/10.3386/w27476

O'Hara, M. (1985). Paired reading in a school for the physically handicapped. *Paired Reading Bulletin*, *1*(1), 16–19. https://files.eric.ed.gov/fulltext/ED285124.pdf

Olkin, I., Dahabreh, I. J., & Trikalinos, T. A. (2012). GOSH–a graphical display of study heterogeneity. *Research Synthesis Methods, 3*(3) 214-223. https://doi.org/10.1002/jrsm.1053

Ottley, P. M. (2003). *Parent teaching and the development of reading skills of at-risk readers* [Doctoral dissertation, University of British Columbia]. https://doi.org/10.14288/1.0091446

Overett, J., & Donald, D. (1998). Paired reading: Effects of a parent involvement programme in a disadvantaged community in South Africa. *British Journal of Educational Psychology*, *68*(3), 347–356. https://doi.org/10.1111/j.2044-8279.1998.tb01296.x

Paige, D. D., Rasinski, T. V., & Magpuri-Lavell, T. (2012). Is Fluent, Expressive Reading Important for High School Readers? *Journal of Adolescent & Adult Literacy*, *56*(1), 67–76. https://doi.org/10.1002/JAAL.00103

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific studies of reading*, *11*(4), 357-383. https://doi.org/10.1080/10888430701530730

Petscher, Y., & Kim, Y.-S. (2011). The utility and accuracy of oral reading fluency score types in predicting reading comprehension. *Journal of School Psychology*, *49*(1), 107–129. https://doi.org/10.1016/j.jsp.2010.09.004

Pigott, T. D., & Polanin, J. R. (2020). Methodological guidance paper: High-quality meta-analysis in a systematic review. *Review of Educational Research*, *90*(1), 24–46. https://doi.org/10.3102/0034654319877153

Pikulski, J. J., & Chard, D. J. (2005). Fluency: Bridge between decoding and reading comprehension. *The Reading Teacher*, *58*(6), 510–519. https://doi.org/10.1598/RT.58.6.2

Poliak, C. P. (1998). *The effect of paired reading on the reading achievement of first-grade students* [Doctoral dissertation, Saint Louis University]. ProQuest Dissertations and Theses Global

Pumfrey, P. D. (1984). Neale analysis of reading ability. In P. Levy & H. Goldstein (Eds.), *Tests in education: Book of critical reviews* (pp. 230–240). Academic Press.

Rasinski, T.V., & Stevenson, B. (2005). The effects of Fast Start Reading: A fluency-based home involvement reading program, on the reading achievement of beginning readers. *Reading Psychology*, *26*(2), 109–125. https://doi.org/10.1080/02702710590930483

Rasinski, T. V. (1994). *Fast start: A parental involvement reading program for primary grade students*. 1–25. https://eric.ed.gov/?id=ED378544

Rasinski, T. V., & Fredericks, A. D. (1991). The Akron Paired Reading Project (working with parents). *Reading Teacher*, *44*(7), 514–515. http://dist.lib.usu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ421254&site=ehost-live

Rasinski, T. V., & Hoffman, J. V. (2003). Oral reading in the school literacy curriculum. *Reading Research Quarterly*, *38*(4), 510–522. https://doi.org/10.1598/RRQ.38.4.5

Rasinski, T. V., Chang, S.-C., Edmondson, E., Nageldinger, J., Nigh, J., Remark, L., Kenney, K. S., Walsh-Moorman, E., Yildirim, K., Nichols, W. D., Paige, D. D., & Rupley, W. H. (2017). Reading Fluency and College Readiness. *Journal of Adolescent & Adult Literacy*, *60*(4), 453–460. https://doi.org/10.1002/jaal.559

Reinking, B., & Bradley, B. (2008). *Formative and design experiments: Approaches to language and literacy research*. Teachers College Press.

Reutzel, D. R., Clark, S. K., Jones, C. D., & Gillam, S. L. (2016). *Young meaning makers—teaching comprehension, Grades K–2*. Teachers College Press.

Richardson, J. R. (1986). Paired reading in Suffolk. *Paired Reading Bulletin*, *2*(1), 97. http://files.eric.ed.gov/fulltext/ED285125.pdf

Ripon, C., Winn, B., & Ingleby, S. (1986). Can parents teach their own children with reading problems? *Paired Reading Bulletin*, *2*(1), 76–81. http://files.eric.ed.gov/fulltext/ED285125.pdf

Ritter, G. W., Barnett, J. H., Denny, G. S., & Albin, G. R. (2009). The effectiveness of volunteer tutoring programs for elementary and middle school students: A meta-analysis. *Review of Educational Research*, *79*(1), 3–38. https://doi.org/10.3102/0034654308325690

Robson, D., Miller, A., & Bushell, R. (1984). The development of paired reading in High Peak and West Derbyshire. *Remedial Education*, *19*(4), 177–183.

Sabatini, J., Wang, Z., & O'Reilly, T. (2019). Relating reading comprehension to oral reading performance in the NAEP fourth-grade special study of oral reading. *Reading Research Quarterly*, *54*(2), 253–271. https://doi.org/10.1002/rrq.226

Samuels, S. J. (1987). Information processing abilities and reading. *Journal of Learning Disabilities*, *20*(1), 18–22. https://doi.org/10.1177/002221948702000104

Schubert, E., Sander, J., Ester, M., Kriegel, H., & Xu, X. (2017). DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN. *ACM Trans Database Systems, 42*(3), 1–19. https://doi.org/10.1145/3068335

Schwanenflugel, P. J., & Benjamin, R. G. (2017). Lexical prosody as an aspect of oral reading fluency. *Reading and Writing*, *30*(1), 143–162. https://doi.org/10.1007/s11145-016-9667-3

Seitz, H. J., & Bartholomew, C. (2014). Reading together: Primary grade children connect with preschoolers. *Young Children*, *69*(1), 22–27. https://search.proquest.com/docview/1511123531/fulltext/88D8240E34744BA7PQ/1?accountid=6724

Shah-Wundenberg, M., Wyse, D., & Chaplain, R. (2013). Parents helping their children learn to read: The effectiveness of paired reading and hearing reading in a developing country context. *Journal of Early Childhood Literacy*, *13*(4), 471–500. https://doi.org/10.1177/1468798412438067

Shanahan, T. (1998). Chapter 6: On the effectiveness and limitations of tutoring in reading. *Review of Research in Education*, *23*(1), 217–234. https://doi.org/10.3102/0091732X023001217

Shany, M. T., & Biemiller, A. (1995). Assisted reading practice: Effects on performance for poor readers in grades 3 and 4. *Reading Research Quarterly*, *30*(3), 382. https://doi.org/10.2307/747622

Simpson, M. (1985). PATCH - Parents and teachers with children at home. *The Paired Reading Bulletin*, *1*(1), 70–71. https://files.eric.ed.gov/fulltext/ED285124.pdf

Skinner, B. F. (1937). Two types of conditioned reflex: A reply Konorski and Miller. *The Journal of General Psychology*, *16*(1), 272–279. https://doi.org/10.1080/00221309.1937.9917951

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, *6*(1), 1–26. https://doi.org/10.1016/j.edurev.2010.07.002

Smith, K. C., Amendum, S. J., & Jang, B. G. (2020). Predicting performance on a 3rd grade high-stakes reading assessment. *Reading & Writing Quarterly*, *36*(4), 365–378. https://doi.org/10.1080/10573569.2019.1649612

Snow, C. E. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. RAND. https://www.rand.org/pubs/monograph_reports/MR1465.html

Sokal, L., Herb, K., Adkins, M., Grills, T., Stewart, C., Priddle, G., Sych-Yereniuk, A., & Chochinov-Harder, L. (2005). Factors affecting inner-city boys' reading: Are male teachers the answer? *Canadian Journal of Urban Research*, *14*(1), 107–130. http://www.jstor.com/stable/44321017

Spalding, B., Drew, R., Ellbeck, J., Livesey, J., Musset, M., & Wales, D. (1984). If you want to improve your reading, ask your Mum': An attempt to involve parents in the reading process at the secondary age level. *Remedial Education*, *19*(4), 157–161.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360–407. https://www.jstor.org/stable/747612

Stothard, S., & Hulme, C. (1991). A note of caution concerning the Neale analysis of reading ability (revised). *British Journal of Educational Psychology*, *61*(2), 226–229. https://doi.org/10.1111/j.2044-8279.1991.tb00978.x

Strong, M. W., & Traynelis-Yurek, E. (1983, March 6). *Behavioral reinforcement within a perceptual-conditioning program of oral reading*. Paper presented at the meeting of the North Carolina Council of the International Reading Association, Winston-Salem, NC.

Teale, W. H., Whittingham, C. E., & Hoffman, E. B. (2020). Early literacy research, 2006–2015: A decade of measured progress. *Journal of Early Childhood Literacy*, *20*(2), 169–222. https://doi.org/10.1177/1468798418754939

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading: A meta-analysis. *Remedial and Special Education*, *25*(4), 252–261. https://doi.org/10.1177/07419325040250040801

Thornton, F. (2012). *The neurological impress method as a reading intervention for students with emotional behavior disabilities* [Doctoral dissertation, Northern Arizona University]. http://search.proquest.com/docview/1282361115/abstract/5BDDA3D8848B4287PQ/1

Tipton, E., Pustejovsky, J. E., & Ahmadi, H. (2019a). A history of meta-regression: Technical, conceptual, and practical developments between 1974 and 2018. *Research Synthesis Methods*, *10*(2), 161–179. https://doi.org/10.1002/jrsm.1338

Tipton, E., Pustejovsky, J. E., & Ahmadi, H. (2019b). Current practices in meta-regression in psychology, education, and medicine. *Research Synthesis Methods*, *10*(2), 180–194. https://doi.org/10.1002/jrsm.1339

Topping, K. J. (1990). *Evaluation of the Kirklees Paired Reading Project* [Doctoral thesis, University of Sheffield]. http://etheses.whiterose.ac.uk/1799/1/DX186618.pdf

Topping, K. J. (1997). A reply to Winter. *Educational Psychology in Practice*, *13*(2), 75–86. https://doi.org/10.1080/0266736970130201

Topping, K. J. (2017). Paired reading and related methods for improving fluency. *International Electronic Journal of Elementary Education*, *7*(1), 57–70. http://www.iejee.com/index.php/IEJEE/article/download/64/62

Topping, K. J., & Bryce, A. (2004). Cross-age peer tutoring of reading and thinking: Influence on thinking skills. *Educational Psychology*, *24*(5), 595–621. https://doi.org/10.1080/0144341042000262935

Topping, K. J, & Lindsay, G. A. (1992a). Paired reading: A review of the literature. *Research Papers in Education*, *7*(3), 199–246. https://doi.org/10.1080/0267152920070302

Topping, K. J., & Lindsay, G. A. (1992b). The structure and development of the paired reading technique. *Journal of Research in Reading*, *15*(2), 120–136. https://doi.org/10.1111/j.1467-9817.1992.tb00028.x

Topping, K. J., Miller, D., Thurston, A., McGavock, K., & Conlin, N. (2011). Peer tutoring in reading in Scotland: Thinking big. *Literacy*, *45*(1), 3–9. https://doi.org/10.1111/ j.1741-4369.2011.00577.x

Topping, K. J., & Whiteley, M. (1993). Sex differences in the effectiveness of peer tutoring. *School Psychology International*, *14*(1), 57–67. https://doi.org/10.1177/0143034393141004

Topping, K. J., Thurston, A., McGavock, K., & Conlin, N. (2012). Outcomes and process in reading tutoring. *Educational Research*, *54*(3), 239–258. https://doi.org/10.1080/00131881.2012.710086

Townsend, J. (1987). Paired reading with parent and peer tutors at High Bank First School – an update. *Paired Reading Bulletin* (3)1, 83-84.

Tymms, P. (1999). *Baseline assessment and monitoring in primary schools: Achievements, attitudes and value-added indicators*. David Fulton.

Tymms, P., Merrell, C., Thurston, A., Andor, J., Topping, K., & Miller, D. (2011). Improving attainment across a whole district: School reform through peer tutoring in a randomized controlled trial. *School Effectiveness and School Improvement*, *22*(3), 265–289. https://doi.org/10.1080/09243453.2011.589859

University of Oregon (2018-2019). 8th Edition of Dynamic Indicators of Basic Early Literacy Skills (DIBELS®): Administration and Scoring Guide. Eugene, OR: University of Oregon. https://dibels.uoregon.edu

Utah State Board of Education (2019). *Partners in Dyad Reading. https://www.schools.utah.gov/file/a63be19c-0988-4dd1-a0bc-ef40b8a85eeb*

Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review*, *23*(1), 87–102. https://doi.org/10.3758/s13423-015-0892-6

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2015). Meta-analysis of multiple outcomes: A multilevel approach. *Behavior Research Methods*, *47*(4), 1274–1294. https://doi.org/10.3758/s13428-014-0527-2

Vaughey, S., & MacDonald, J. (1986). Paired reading project with Asian families. *Paired Reading Bulletin*, *2*(1), 6–9. http://files.eric.ed.gov/fulltext/ED285125.pdf

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3). https://doi.org/10.18637/jss.v036.i03

Vo, T. N. (2011). *Paired reading intervention for limited English proficient students* [Doctoral thesis, University of Dayton]. https://etd.ohiolink.edu/pg_10?0::NO:10:P10_ACCESSION_NUM:dayton1311774375

Weiss, J. A., Thurlow, M. L., Christenson, S. L., & Ysseldyke, J. E. (1989). *Paired reading with adult volunteer tutors as a reading intervention for students with reading difficulties*. https://eric.ed.gov/?id=ED305606

Welsh, M., & Roffe, M. (1985). Paired reading projects with Asian families. *Paired Reading Bulletin*, *1*(1), 34–37. http://files.eric.ed.gov/fulltext/ED285124.pdf

Whithear, J. (2011). A review of fluency research and practices for struggling readers in secondary school. *Literacy Learning: The Middle Years*, *19*(1), 18–28. http://dist.lib.usu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=58027638&site=ehost-live

Wilfong, L. G. (2008). Building fluency, word-recognition ability, and confidence in struggling readers: The poetry academy. *The Reading Teacher*, *62*(1), 4–13. https://doi.org/10.1598/RT.62.1.1

Winter, S. (1988). Paired reading: A study of process and outcome. *Educational Psychology*, *8*(3), 135–151. https://doi.org/10.1080/0144341880080301

Winter, S. (1997). Lessons from paired reading: A reply to Topping. *Educational Psychology in Practice*, *13*(2), 87–95. https://doi.org/10.1080/0266736970130202

Young, C., Durham, P., & Rosenbaum-Martinez, C. (2018). A stacked approach to reading intervention: Increasing 2nd- and 3rd- graders' independent reading levels with an intervention program. *Journal of Research in Childhood Education*, *32*(2), 181–189. https://doi.org/10.1080/02568543.2017.1418771

Young, C., Lagrone, S., & McCauley, J. (2020). Read like me: An intervention for struggling readers. *Education Sciences*, *10*(3), 57. https://doi.org/10.3390/educsci10030057

Young, C., Mohr, K. A. J., & Rasinski, T. (2015). Reading together: A successful reading fluency intervention. *Literacy Research and Instruction*, *54*(1), 67–81. https://doi.org/10.1080/19388071.2014.976678

Young, C., Pearce, D., Gomez, J., Christensen, R., Pletcher, B., & Fleming, K. (2018). Read Two Impress and the neurological impress method: Effects on elementary students' reading fluency, comprehension, and attitude. *The Journal of Educational Research*, *111*(6), 657–665. https://doi.org/10.1080/00220671.2017.1393650

Young, C., Rasinski, T., & Mohr, K. A. J. (2016). Read Two Impress: An intervention for disfluent readers. *Reading Teacher*, *69*(6), 633–636. https://doi.org/10.1002/trtr.1391

Young, C., & Rasinski, T. V. (2017). *Tiered fluency instruction: Supporting diverse learners in grades 2-5*. Capstone.

Zutell, J., & Rasinski, T. V. (1991). Training teachers to attend to their students' oral reading fluency. *Theory Into Practice*, *30*(3), 211–217. https://doi.org/10.1080/00405849109543502

APPENDIX

CODING MANUAL

Coding Manual

This coding manual was used to calculate an effect size for each fluency and comprehension related outcome, and record details on potentially significant moderators. Spreadsheet software was used to record this information. Critically, this coding manual is organized by effect size rather than study. Each independent effect size occupied a single row of the spreadsheet and studies with multiple effect sizes occupied multiple rows. Organizing the coding manual by effect size, rather than study, accounts for the multilevel analysis. The three-level analysis modeled effect sizes nested within covariance between effect sizes, nested within studies. Any data not reported in an extant study was left blank in the coding manual. The remainder of this appendix indicates the layout of the spreadsheet, and details regarding the gathered information.

Coding Description and Instructions

Study Details

- Column A: Study Authors
    - Authors of the study, recorded in APA (7th ed.) format
- Column B: Study Synopsis
    - Brief summary of study goals, measures, and participants. This information informed tables in the results section.

Hierarchy of Statistical Model

- Column C: Study Number
    - Each study coded with a separate number. These data formed Level 3 of the multilevel model.
- Column D: Effect Number
    - Each effect size coded with a separate number. These data allow for a multivariate analysis (multiple effects within a single study), and formed Level 2 of the multilevel model.
- Column E: Effect Size
    - Effect size, calculated as Hedge's *g*. These data formed Level 1 of the multilevel analysis.

Variables for Moderator Analysis

- Column F: Reading Construct of Effect Size
    - Reading construct coded as fluency, comprehension. These data were used in the moderator analysis. The following indicates the coding process for each category:
        - Effect sizes coded as fluency met the following inclusion criteria:
            - Measures of rate, such as words correct per minute (WCPM)
            - Measures of accuracy for assessments that use connected text
            - Measures of prosody, such as the Multidimensional Fluency Scale (MDFS; Raskinski & Zutell, 1991)
        - Effect sizes coded as comprehension met the following inclusion criteria:
            - Measures of reading comprehension using connected text
- Column G: Assessment
    - Name of assessment used.
- Column H: Year
    - Year of publication, these data were used in the moderator analysis.
- Column I: Publication Type
    - Format of publication, coded as peer-reviewed journal, dissertation, conference report, or other. Details of any studies coded as 'other' were recorded separately. These data were used in the moderator analysis.
- Column J: Study Design
    - Study design was coded using the *What Works Clearinghouse Group Design Standards* (What Works Clearinghouse [WWC], 2020.). Studies were coded as random, non-random, or single group design. These data were dummy coded as categorical variables, and used in the moderator analysis. Each study received a code corresponding with the following three designations.
        - *Meets WWC Group Design Standards Without Reservation (2)*

        - Studies assigned this code used a random assignment to condition and evidence low sample attrition.

    - *Meets WWC Group Design Standard With Reservation (1)*

        - Studies assigned this code used a non-random condition assignment, and equivalence established at baseline between conditions.

    - *Does Not Meet WWC Group Design Standards (0)*

        - Studies assigned this code used a non-random condition assignment, and non-equivalence at baseline. Designs that omitted a control group or used a single group design were also be assigned this code.

- Column K: Age of Participants

    - Age of participants from the study. Studies that indicate a range of ages were calculated as an average. These data were used in the moderator analysis.

- Column L: Tutoring Method

    - Method of Paired Oral Reading used in the study. Methods were coded as Neurological Impress Method, Paired Reading, Dyad Reading, or Read Two Impress. These data were used in the moderator analysis.

- Column M: Tutor Type

    - Type of tutor used, coded as adult, cross-age peer, or same-age peer. These data were used in the moderator analysis.

- Column N: Minutes

    - Total dosage of intervention coded as minutes. This column was left blank for studies that do not report the information or allow for a calculation. These data were used in the moderator analysis.

- Column O: Weeks

    - Total duration of intervention coded as weeks. This column was left blank for studies that do not report the information or allow for a calculation. These data were used in the moderator analysis.

Data Used to Calculate Effect Sizes

- Column P: Page Number of Effect Size Data
    - Page numbers(s) where the effect size, or data used to calculate an effect size, is listed. This column information was included for reference purposes and was not used in the multilevel analysis.
- Column Q: Treatment Sample
    - Sample size of the treatment group.
- Column R: Control Sample
    - Sample size of the control group.
- Column S: Pooled Sample Size
    - Combined treatment and control sample size.
- Column T: Control Group Pretest Mean
    - Mean of the control group at pretest.
- Column U: Control Group Pretest Standard Deviation
    - Standard deviation of the control group pretest mean.
- Column V: Control Group Posttest Mean
    - Mean of the control group at posttest
- Column W: Control Posttest Standard Deviation
    - Standard deviation of the control group posttest mean.
- Column X: Treatment Group Pretest Mean
    - Mean of the treatment group at pretest.
- Column Y: Treatment Group Pretest Standard Deviation
    - Standard deviation of the treatment group pretest mean.
- Column Z: Treatment Group Posttest Mean
    - Mean of the treatment group at posttest

- Column AA: Treatment Posttest Standard Deviation
    - Standard deviation of the treatment group posttest mean.

CURRICULUM VITAE

JACOB D. DOWNS

Elementary Literacy Coordinator
Cache County School District

Adjunct Instructor
Utah State University
School of Teacher Education and Leadership

Phone: (435) 232-4020
Email: Jake.Downs@usu.edu

Education

PhD — May 2021
Curriculum & Instruction, Literacy and Leadership Emphasis
Utah State University
Dissertation: A Multilevel Meta-Analysis of Paired Oral Reading Methods in Elementary Classrooms

MLTID — May 2017
Master of Learning Technology and Instructional Design
Utah State University

BA — May 2015
Bachelor of Arts in Elementary Education
Utah State University
1-6 Grade License, State of Utah

Employment History

Elementary Literacy Coordinator (June 2021-Present)
Cache County School District
Major responsibilities include the coordination, articulation, and alignment of all elementary literacy instruction and curriculum within Cache County School District. Included in this responsibility is reviewing and analyzing student achievement data to identify specific areas for ongoing professional learning. Also included is oversight of student support using a Response to Intervention (RTI) framework.

Instructional Coach (Aug 2020-May 2021)
Cache County School District
Major responsibility consists of training and supporting all novice third and fourth grade

teachers within Cache County School District. Support is also given to veteran teachers who have transferred from another school district. Intake support includes weekly trainings for the first two months of the school year. These trainings orient teachers to evidence-based, high-leverage practices in classroom management, literacy, and other content areas. Ongoing support consists of following the Impact Cycle (Knight, 2017) model of classroom instructional coaching. Coach and candidate collect data in key areas, select a high-leverage goal, and engage in co-teaching throughout a three-week cycle. Additional data is collected throughout and at the end of the cycle to ascertain progress toward the targeted goal. Other responsibilities include developing grade-level, school, and district wide professional development in literacy, service on various committees intended to support classroom instruction, and serving as a liaison between classroom teachers and district entities.

4th Grade Teacher, (Aug 2015-June 2020)
White Pine Elementary, Cache County School District
Responsibilities included teaching the 4th grade Utah Core across all content areas to students of varying abilities. Populations served include students with Individualized Education Plans, students with Section 504 plans, English language learners, gifted and talented students, and students who experienced trauma. Extra-curricular activities included serving as the chair of our Student and Teacher Assistance Team (STAT), safety committee member, robotics club advisor, hiring committee, and school ski day liaison/coordinator.

2nd Grade Teacher, (April 2015-June 2015)
Summit Elementary, Cache County School District
Completed the school year for a teacher who retired in early spring. Responsibilities included planning and implementing curriculum in all content areas aligned to the Utah Core for students of varying abilities.

AmeriCorps Member (Aug 2012- June 2014)
Birch Creek Elementary, Cache County School District
Responsibilities included coordinating Read Today tutoring program, recruiting and training volunteers from the community to deliver targeted reading intervention to at-risk students, delivering reading instruction to at-risk first and third graders in a small group intervention, and serving as liaison between paraprofessionals and school administration.

Licensure and Certifications

Level 2 Professional Educator License, State of Utah
Educational Technology Endorsement, State of Utah

Academic Contributions

Research Interests

- Understanding how methods of paired oral reading influence fluency and comprehension outcomes for elementary-age students
- Investigating how models of comprehension, such as Kintsch's Construction Integration Model, are effectively integrating into classroom instruction
- Understanding how differences between oral reading fluency and reading comprehension influence classroom instructional practices in the elementary years
- Investigating best practices in teaching reading comprehension in the upper elementary grades.

Academic Publications

Mohr, K. A. J., Chou, P., Tsai, Z., & Downs, J. D. (Manuscript Under Review). Using Read-and-Think-Alouds to Teacher Inferencing: Supporting Comprehension From the Beginning

Downs, J.D. (Manuscript Accepted). Upper Elementary Reading Comprehension: The Construction-Integration Model of Reading and Transactional Strategies Instruction. Utah Journal of Literacy.

Mohr, K. A. J., Downs, J. D., & Mohr, E. S. (2020). Mindful Reading: Eye-Tracking Evidence for Goal-Directed Instruction. Journal of Adolescent & Adult Literacy, 00(00). https://doi.org/10.1002/jaal.1099

Downs, J. D., Mohr, K. A. J., & Barrett, T. S. (2020). Determining the academic and affective outcomes of dyad reading among third graders. The Journal of Educational Research, 113(2), 120–132. https://doi.org/10.1080/00220671.2020.1755615

Academic Contributions

Member, Practitioner Review Board, The Reading Teacher (March 2021- Present). Role includes anonymously peer-reviewing articles submitted to *The Reading Teacher.* Major responsibilities include providing feedback for authors on ways to make submitted articles more valuable for classroom teachers, and evaluating articles for alignment with current research evidence.

Academic Presentations

Downs, J.D. (Accepted) "Paired Oral Reading: A Multilevel Meta-Analysis of Related Practices" Virtual Poster Session at the 28th annual conference of the Society for the Scientific Study of Reading Conference, Online, July 13-16, 2021

Downs, J.D. "Acquiring Academic Skill: Breadth and Depth" Invited presentation to TEAL 7321 Foundations of Language and Literacy Course, Utah State University, June 21, 2021

Downs, J.D., & Mohr, K.A. (Accepted, Conference Cancelled due to COVID-19) "The Influences of Dyad Reading on the Reading Attitudes of Third Graders" Presentation at the 27th annual conference of the Society for the Scientific Study of Reading Conference in Newport Beach, July 8-11, 2020

University Courses Taught

ELED 3100- Classroom Reading Instruction (Fall 2019, Spring 2020, Fall 2020, Spring 2021, Scheduled Fall 2021)
EDUC 3660- Educational Psychology (Spring 2019)
TEAL 3000- Social, Historical, and Cultural Foundations (Fall 2018)
TEAL 5560- Elementary Education Practicum Experience (Fall 2018)
TEAL 5500- Integration and Innovation of Technology (Summer 2017-Spring 2018)

Course Development

ELED 3100 Elementary Classroom Reading Instruction (Summer 2019, Summer 2020, Summer 2021)
Collaborated with Dr. Cindy Jones and other course instructors to better align coursework with effective classroom methods of literacy instruction. Major responsibilities included reworking the fluency and vocabulary lesson plan assignments to incorporate recommendations from IES Practice Guides and other research-based evidence. Minor responsibilities included adjusting course schedule and attending to weekly alignment with course objectives.

TEAL 3000 Social, Historical, and Cultural Foundations in Education (Summer, 2018)
Collaborated with Dr. Emma Mecham in course adjustments. Main responsibility included course design on the LMS Canvas for a better student experience. Minor responsibilities included contributing to course content, assignments, and overall course structure.

Professional Presentations

Downs, J.D., Covili, J., Sloan, D., Lara, F., Vendeventer, T., & Martin, J. "Utah Podcasters Unite! A Panel Discussion Featuring Utah Education Podcasts." Presentation at the annual Utah Coalition for Educational Technology conference in Provo, Utah March 11, 2020.

Downs, J.D. "Students Mastering Their Own Data with Google Sheets." Presentation at the annual Utah Coalition for Educational Technology Conference in Salt Lake City, Utah March 22, 2018.

Professional Contributions

Chair, Student and Teacher Assistance Team (STAT), (August 2016-June 2020)
Responsibilities include coordinating intervention team for at-risk students at a school of 400 students. The committee supports teachers with skill development, intervention assistance, and data analysis. Populations served include students with an Individualized Education Plan (IEP), students with a Section 504 Plan, English language learners, students with PTSD, students who are at risk due to poverty, students with emotional disturbances, students with severe anxiety, and students experiencing unexpected underachievement.

Member, White Pine School Safety Committee, (December 2019- June 2020)
Responsibilities include planning various aspects of school safety in preparation for future emergencies. Extra emphasis placed on plan for reuniting students with parents after an emergency has occurred as well as training faculty and staff in school emergency procedures.

Member, 4th Grade ELA-vate Utah Committee (December 2017- Aug 2018)
Responsibilities included contributing to revising components of the ELA-vate Utah 4th Grade Language Arts Curriculum and supporting teachers with added developments.

Professional Awards

Recipient, $1000 Scholarship, Utah Retired School Educators Association, 2019
Teacher of the Year, White Pine Elementary, 2018-2019 School Year
Hats Off Award, Cache Education Foundation, 2017-2018 School Year

Professional and Academic Service

Downs, J. (Host and Producer, Teaching Literacy Podcast). Major focus is on bridging literacy research into practice. Interviews are conducted with literacy researchers about their work and how to adapt it into classroom instruction.

Parsons, S. (Presenter) & Downs, J. (Host/Producer). (2021, June 4) *Principles of Effective Literacy Instruction with Dr. Seth Parsons* [Audio Podcast]. Retrieved from https://teachingliteracypodcast.com/e24-principles-of-effective-literacy-instruction-with-dr-seth-parsons/

Mohr, K. (Presenter), Mohr, E. (Presenter), & Downs, J. (Host/Producer). (2021, June 4) *Eye Tracking Evidence for Goal Directed reading With Dr. Kit Mohr and Dr.*

Eric Mohr [Audio Podcast]. https://teachingliteracypodcast.com/e23-eye-tracking-evidence-for-goal-directed-reading-with-dr-kit-mohr-and-dr-eric-mohr/

Clark, S.C. (Presenter) & Downs, J. (Host/Producer). (2021, May 13) *Integrating Literacy and Science Instruction for Early Elementary Students With Dr. Sarah K. Clark* [Audio Podcast]. Retrieved from https://teachingliteracypodcast.com/e-22-integrating-literacy-and-science-instruction-for-early-elementary-students-with-dr-sarah-k-clark/

Snow, P. (Presenter) & Downs, J. (Host/Producer). (2020, August 19) *SOLAR: Science of Language and Reading with Dr. Pamela Snow* [Audio Podcast]. Retrieved from https://teachingliteracypodcast.com/e-21-solar-science-of-language-and-reading-with-dr-pamela-c-snow/

Shanahan, T. (Presenter) & Downs, J. (Host/Producer). (2020, August 5) *20 Years of NRP 2000 with Dr. Tim Shanahan* [Audio Podcast]. Retrieved from https://teachingliteracypodcast.com/e20-20-years-of-nrp-2000-with-dr-tim-shanahan/

Mohr, K. (Presenter) & Downs, J. (Host/Producer). (2020, July 27*). Academic and Attitude Outcomes With Third Graders With Dr. Kit Mohr and Jake Downs* [Audio Podcast]. Retrieved from https://teachingliteracypodcast.com/e19-dyad-reading-academic-and-attitude-outcomes-with-third-graders/

Baker, D. (Presenter), Santoro L. (Presenter), & Downs, J. (Host/Producer). (2020, July 20*). Read Alouds Using a Structural Layering Approach* [Audio Podcast]. Retrieved from https://teachingliteracypodcast.com/e18-read-alouds-using-a-structural-layering-approach/

Patrick S. (Presenter) & Downs, J. (Host/Producer). (2020, June 28) *Homogenous Grouping with Dr. Susan Patrick* [Audio Podcast]. Retrieved from https://teachingliteracypodcast.com/e17-homogeneous-grouping-with-dr-susan-patrick/

Delacruz, S. (Presenter), & Downs, J. (Host/Producer). (2020, June 9) *Supporting ELA Instruction Through Coding with Dr. Stacy Delacruz*. [Audio Podcast]. Retrieved from https://teachingliteracypodcast.com/e16-supporting-ela-instruction-through-coding-with-dr-stacy-delacruz/

Millier, M. (Presenter), & Downs, J. (Host/Producer). (2020, May 16) *Remote Learning 101 with Matt Miller* [Audio Podcast] Retrieved from https://teachingliteracypodcast.com/episode-15-remote-learning-101-with-matt-miller/

Tollefson D. (Presenter. & Tollefson C. (Presenter), & Downs, J. (Host/Producer). (April 26, 2020) *Resilient Students: Bridging Body and Mind with Dr. Derrik and Cassie Tollefson.* [Audio podcast]. Retrieved from https://teachingliteracypodcast.com/episode-14-resilient-students-bridging-body-and-mind-with-dr-derrik-and-cassie-tollefson/

Hiebert, E. (Presenter), & Downs, J. (Host/Producer). (2020, April 10). *How Words Work with Dr. Elfrieda Hiebert.* [Audio podcast]. Retrieved from https://teachingliteracypodcast.com/episode-13-how-words-work-with-dr-elfrieda-hiebert/

Hayden, E. (Presenter), & Downs, J. (Host/Producer). (2020, March 14). *Silent reading rate and comprehension with Dr. Emily Hayden.* [Audio podcast]. Retrieved from https://teachingliteracypodcast.com/episode-12-silent-reading-rate-and-comprehension-with-dr-emily-hayden/

Phillips Galloway, E. (Presenter) & Downs, J. (Host/Producer). (2020, Februrary 29). *Beyond reading comprehension: core academic language skills with Dr. Emily Phillips Galloway.* [Audio podcast]. Retrieved from https://teachingliteracypodcast.com/episode-11-beyond-reading-comprehension-core-academic-language-skills-with-dr-emily-phillips-galloway/

Hoffman, E.B. (Presenter), Whittingham, C.E. (Presenter), & Downs, J. (Host/Producer). (2020, February 4). *Reviewing a decade of early literacy research with Drs. Emily Brown Hoffman and Colleen Whittingham.* [Audio podcast]. Retrieved from https://teachingliteracypodcast.com/episode-10-reviewing-a-decade-of-early-literacy-research-with-dr-emily-brown-hoffman-and-colleen-whittingham/

Vaughn, M (Presenter) & Downs, J. (Host/Producer). (2020, January 9). *Student agency in literacy with Dr. Margaret Vaughn.* [Audio podcast]. Retrieved from https://teachingliteracypodcast.com/episode-09-student-agency-in-literacy-with-dr-margaret-vaughn/

Vaughn, M (Presenter) & Downs, J. (Host/Producer). (2019, December 12). *Adaptive literacy teaching with Dr. Margaret Vaughn.* [Audio podcast]. Retrieved from https://teachingliteracypodcast.com/episode-08-adaptive-literacy-teaching-with-dr-margaret-vaughn/

Zhang, C (Presenter) & Downs, J (Host/Producer). (2019, November 30). *Integrating writing with morning meeting for preschoolers with Dr. Chenyi Zhang.* [Audio podcast]. Retrieved from https://teachingliteracypodcast.com/episode-07-integrating-writing-with-morning-meeting-for-preschoolers-with-dr-chenyi-zhang/

Reutzel, D.R. (Presenter) & Downs, J. (Host/Producer). (2019, November 24). *Comprehension: CI theory and practice part two with Dr. D. Ray Reutzel.* [Audio podcast]. Retrieved from https://teachingliteracypodcast.com/episode-06-comprehension-ci-theory-and-practice-part-2-with-dr-d-ray-reutzel/

Reutzel, D.R. (Presenter) & Downs, J. (Host/Producer). (2019, November 16). *Comprehension: CI theory and practice with Dr. D. Ray Reutzel.* [Audio podcast]. Retrieved from https://teachingliteracypodcast.com/episode-05-comprehension-ci-theory-and-practice-with-dr-d-ray-reutzel/

Young, C. (Presenter) & Downs, J. (Host/Producer). (2019, November 9). *Tiered fluency support part two.* [Audio podcast]. Retrieved from https://teachingliteracypodcast.com/episode-04-tiered-fluency-instruction-part-2/

Young, C. (Presenter) & Downs, J. (Host/Producer). (2019, November 2). *Tiered fluency Support Part One.* [Audio podcast]. Retrieved from https://teachingliteracypodcast.com/episode-03-fluency-instruction-part-1/

Strong, J.Z. (Presenter) & Downs, J. (Host/Producer). (2019, October 26). *Supporting Students in Difficult Text.* [Audio podcast]. Retrieved from https://teachingliteracypodcast.com/episode-02-supporting-students-in-difficult-text/.

Moses, L. (Presenter) & Downs, J. (Host/Producer). (2019, October 16). *Behavior of First Graders During Independent Reading Time.* [Audio podcast]. Retrieved from https://teachingliteracypodcast.com/episode-01-behavior-of-first-graders-during-independent-reading-time/.