ABSTRACT

Phonemic Awareness Instruction with Children at Risk of Reading Failure

David M. Rehfeld, Ph.D.

Mentor: Tracey N. Sulak, Ph.D.

Although reading is a set of skills critical to long term educational and vocational outcomes, many children in the United States are at risk of reading failure for a variety of reasons. For these children, establishing sufficient levels of phonemic awareness in the early grades is critical for the successful development of word reading skills and, indirectly, reading comprehension. The present dissertation combines the results of a systematic review and meta-analysis of the existing literature on the use of phonemic awareness interventions with children at risk of reading failure with a single case investigation of one such intervention used with second grade students struggling to read. Based on a review of the existing literature, the effects of phonemic awareness interventions used with at-risk children are significant but the magnitude of these effects vary with respect to the target outcome, with average Hedge's *g* values ranging from .25 to .57. The results continue to indicate that phonemic awareness instruction is generally the most effective when graphemes are incorporated appropriately and intervention is provided sooner rather than later. Instruction can be effective when provided by a variety of school personnel, though the strongest outcomes were produced by interventions

implemented by speech-language pathologists. Intervention has also been demonstrated to be efficacious when delivered individually, in small groups, or in larger groups such as through whole class instruction. The adjoining single case investigation of contextualized phonemic awareness instruction provided to second grade students also indicates that phonemic awareness instruction conducted over a relatively short period of time can affect significant substantial change. Based on the results of the included single case research, contextualizing phonemic awareness instruction might help affect generalized change across multiple outcomes related to children's ability to successfully read. Together, this meta-analysis and single case investigation continue to support the provision of appropriately designed phonemic awareness instruction children at risk of reading failure to support the acquisition of basic reading skills necessary to engage in the general education curriculum.

Phonemic Awareness Instruction with Children at Risk of Reading Failure

by

David M. Rehfeld, B.A., M.Ed., M.S.

A Dissertation

Approved by the Department of Educational Psychology

____

Grant Morgan, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of

Doctor of Philosophy

Approved by the Dissertation Committee

____

Tracey Sulak, Ph.D., Chairperson

____

Janet Bagby, Ph.D.

____

Nicholas Benson, Ph.D.

____

Stephanie Gerow, Ph.D.

____

Diane Loeb, Ph.D.

Accepted by the Graduate School
August 2021

____

J. Larry Lyon, Ph.D., Dean

Page bearing signatures is kept on file in the Graduate School.

TABLE OF CONTENTS

LIST OF FIGURES v
LIST OF TABLES vi
ACKNOWLEDGMENTS viii
CHAPTER ONE 1
  Introduction 1
    Reading Achievement and Phonological Awareness 1
CHAPTER TWO 4
  Literature Review 4
    Phonemic Awareness 4
    Intervention 6
CHAPTER THREE 10
  Phonemic Awareness Instruction with At-Risk Readers: A Meta-Analysis 10
    Method 12
    Results 25
    Discussion 57
CHAPTER FOUR 66
  Contextualized Phonemic Awareness Instruction with At-Risk Readers 66
    Method 69
    Results 91
    Discussion 103
CHAPTER FIVE 109
  Conclusion 109
APPENDIX 115
  List of Studies Included in Meta-Analysis 115
REFERENCES 130

LIST OF FIGURES

Figure 3.1. PRISMA flow diagram 35

Figure 3.2. Boxplots of phonemic awareness outcomes 45

Figure 3.3. Forest plot for studies reporting phonological awareness composite scores...47

Figure 3.4. Forest plot for studies reporting segmentation scores....................................52

Figure 3.5. Forest plot for studies reporting blending scores 56

Figure 3.6. Forest plot for studies reporting deletion scores 60

Figure 3.7. Forest plot for studies reporting FSID scores 63

Figure 4.1. Probe accuracy percentages by session for participant one 104

Figure 4.2. Probe accuracy percentages by session for participant two 105

Figure 4.3. Probe accuracy percentages by session for participant three 106

Figure 4.4. Probe accuracy percentages by session for participant four 107

LIST OF TABLES

Table 3.1. Study quality summary statistics using Troia’s (1999) weighted scores 29

Table 3.2. Studies included for review 37

Table 3.3. Summary of overall study quality 42

Table 3.4. Subgroup analyses of phonemic awareness composite outcomes 49

Table 3.5. Outcomes by cumulative intervention intensity percentile rank 50

Table 3.6. Subgroup analyses of segmentation scores 53

Table 3.7. Subgroup analyses of blending scores 57

Table 3.8. Subgroup analyses of deletion scores 61

Table 3.9. Subgroup analyses of FSID scores 64

Table 3.10. Outcomes of tests for small study effects 66

Table 3.11. Summary of intervention effects by phonemic awareness outcome 67

Table 4.1. STAR Early Literacy domains and associated skills 81

Table 4.2. CTOPP-2 composite and subtest descriptions 84

Table 4.3. DIBELS-8 scoring procedures 86

Table 4.4. KBIT-2 scores by participant 87

Table 4.5. CTOPP-2 pretest scores by participant 88

Table 4.6. DIBELS-8 pretest scores by participant 89

Table 4.7. DIBELS-8 pretest risk classifications by participant 89

Table 4.8. Platform books utilized in intervention 95

Table 4.9. Prompting hierarchy 96

Table 4.10. Interobserver agreement data 98

Table 4.11. Treatment fidelity checklist 100

Table 4.12. Treatment fidelity session counts 100

Table 4.13. Treatment fidelity data 101

Table 4.14. Dosage summary statistics by participant 103

Table 4.15. Baseline and intervention probe data by participant 108

Table 4.16. CTOPP-2 posttest scores by participant 109

Table 4.17. DIBELS-8 pretest, posttest, and gain scores by participant 109

Table 4.18. DIBELS-8 posttest risk classifications by participant 110

Table 4.19. Effect estimates by participant 111

Table 4.20. Baseline trend statistics by participant 112

ACKNOWLEDGMENTS

I would like to thank Dr. Tracey Sulak not only for her guidance in developing this series of projects but also her willingness to invest in and mentor me over the past few years. I would also like to thank the rest of my committee members, Drs. Janet Bagby, Nicholas Benson, Stephanie Gerow, and Diane Loeb for their time, effort, and grace in refining this project, especially their guidance on conducting this project safely and effectively during the covid-19 pandemic. Additionally, this project would not have been possible without the help of my friends and colleagues: Marie Kirkpatrick, Nicole O'Guinn, and Noah Padgett, whose professional and personal support were integral to its completion. On a more personal level, I would not have been able to complete this project without the help of my University Baptist Church community—especially Austin and Nikki Lemmons, Marcus Mataga, Jill Riley and the rest of my small group over the past four years—who have helped me to understand the importance of intersectionality and the human condition. Finally, an infinite number of words would never be enough to thank my family for their unwavering love and support: here, the fewest words acknowledge the greatest debt I can never repay.

CHAPTER ONE

Introduction

Phonemic awareness, or the ability to attend to and manipulate individual sounds within spoken words, has long been recognized as contributing to the development of reading in children (Liberman et al., 1974). More specifically, it has been strongly linked to word decoding skills and thus is indirectly related to long-term success in reading comprehension (Ehri et al., 2001; Kamhi & Catts, 2012). Beginning in approximately third grade, children are expected to learn information from sources other than verbal instruction (e.g., textbooks) and then demonstrate understanding of what they have read—this is commonly referred to as "reading to learn". Children who fail to develop adequate phonemic awareness skills necessary for the acquisition of fluent word decoding will have to allocate more cognitive resources to identifying the words in a text before they can even attempt to understand its meaning. Fortunately, there is substantial research indicating that interventions and instruction designed to develop phonemic awareness in children can be effective (e.g., Ehri et al., 2001; Suggate, 2016). Much of this research, however, is focused on the decontextualized instruction of these skills and does not situate them within the natural context for which they are most important: reading.

Reading Achievement and Phonological Awareness

The Simple View of reading formulates reading comprehension as a product of word decoding skills and listening comprehension (Catts, 2009; Gough & Tunmer, 1986).

Further research into the development of reading comprehension has indicated that one of the best predictors of reading comprehension is an individual's familiarity with content knowledge (Kamhi & Catts, 2017). Accordingly, literacy activities for beginning readers should focus on building their word decoding skills using familiar content and/or highly preferred topics, allowing them to focus on decoding the words accurately and developing their alphabetic insight or the understanding that specific letters are associated with specific sounds (Adams, 1990; Kamhi, 2002). In order to develop word decoding skills, though, children must first understand that words themselves are not a wholly indestructible unit. They must develop some understanding that words can be broken down into smaller parts and also built from those same components—this is what researchers call phonological awareness.

"Phonological awareness refers to an individual's awareness of the sound (phonological) structure of spoken words" (McNeill et al., 2017, p. 302). That is, phonological awareness is a broad, metalinguistic skill that represents an individual's understanding that words can be broken down into smaller units. Phonological awareness encompasses a wide variety of skills, ranging from recognizing rhyme (i.e., some words sound like others) to substituting sounds within words to change their meaning (e.g., substituting the first sound in mad to change it to bad). It requires an individual to be aware of both a word's overall syllabic structure as well as its intersyllabic structure (i.e., the structure within syllables).

Broadly speaking, phonological awareness is a strong predictor of long-term reading achievement (Hogan et al., 2005; Torgesen et al., 1994). Before children enter school, during the period many scholars refer to as the emergent literacy period, critical skills are developed that set them up for success in learning to read. The main skills of interest related to reading for children entering kindergarten are their print knowledge and their phonological awareness, both of which can be developed through joint book reading (Bus et al., 1995; Justice & Ezell, 2004). Because this period precedes formal schooling, however, the amount of literacy knowledge with which children enter kindergarten or first grade is highly variable (Kamhi & Catts, 2012; Tiruchittampalam et al., 2018). While joint book reading is helpful in developing both print knowledge and phonological awareness, phonological awareness can also be developed through activities such as nursery rhymes that are not necessarily tied to printed text (Piasta et al., 2012). This difference makes phonological awareness a less culturally biased predictor of reading achievement than print knowledge, especially for children from low socioeconomic backgrounds or others with limited engagement in literacy activities before formal schooling (Kamhi & Catts, 2012).

As children develop, they learn that words can be broken down into syllables and, as they engage in activities related to word structure such as nursery rhymes, they begin to learn that these syllables themselves can be broken down into smaller onset-rime units (McNeill et al., 2017). The onset of a syllable is the consonant or consonant cluster/blend that precedes the vowel, while the rime of a syllable is the vowel plus any consonants that follow it. As children further develop their understanding of these concepts—regardless of whether they know the terminology—they begin to develop their awareness that the onsets and rimes themselves can be further broken down into individual speech sounds (i.e., phonemes). This awareness that words can be broken down into combinations of individual speech sounds is termed *phonemic* awareness.

CHAPTER TWO

Literature Review

Although phonological awareness encompasses a wide variety of skills and is strongly linked to reading achievement (Hogan et al., 2005), previous research has indicated that *phonemic* awareness is the better predictor of long-term reading achievement (Ehri et al., 2001; Melby-Lervåg et al., 2012; Suggate, 2016). The difference between phonological awareness and phonemic awareness is primarily one of depth, with phonological awareness at the word and syllable levels being characterized as "shallow" and phonological awareness at the single sound level being characterized as "deep" (Schuele & Boudreau, 2008).

Phonemic Awareness

Terminologically, phonemic awareness refers to both a construct and a set of skills. As a construct, phonemic awareness refers to an individual's understanding that spoken words are comprised of individual speech sounds referred to as phonemes, the smallest meaningful unit of speech within a given language. Phonemic awareness also, however, refers to a distinct set of skills involving those phonemes, such as segmenting words into their constituent phonemes (e.g., breaking "cat" down into /k/ /ae/ /t/), blending phonemes into words (e.g., building /k/ /ae/ /t/ into "cat"), and otherwise manipulating phonemes within words (e.g., substituting /m/ for the /k/ to change "cat" into "mat"). Kamhi & Catts (2012) provided a definition for phonemic awareness that attempts to recognize that it is simultaneously a construct and a set of skills, saying that it involves a more or less explicit understanding that words are composed of sound segments smaller than a syllable, as well as knowledge, or awareness, of the distinctive features of individual phonemes themselves. It is this latter knowledge of the identity of individual phonemes themselves that continues to increase after an initial understanding of the phonemic structure of words is acquired.

Previous research has found that children with poor phonemic awareness in the early elementary grades often continue to lag behind their typically developing peers for the remainder of their early educational career, if not longer (Torgesen et al., 1994). Without intervention, these children will struggle to demonstrate acceptable academic progress as the curriculum becomes more difficult and relies more heavily information gained from independent reading assignments. Phonemic awareness is also one of the strongest predictors of reading achievement *before children even begin to read* (e.g., Fletcher et al., 1994), which makes it a valuable measure for early childhood educators working to identify those children at risk of reading failure for whom intervention is desperately needed.

Phonemic awareness, however, is not a perfect predictor of reading success. In some cases, phonemic awareness screenings can result in unacceptably high false positive rates, meaning that children who are predicted to struggle with reading actually do fine with minimal or no intervention (Alonzo et al., 2020; Torgesen et al., 2003). For example, while there is a strong relationship between developmental language disorder and the later development of reading disorders (Catts et al., 2005), many children with developmental language disorder present with below average phonemic awareness skills and go on to be successful readers (Alonzo et al., 2020). Such difficulties in the accurate identification of children with reading disorders and difficulties based on phonemic awareness screenings mean that phonemic awareness assessments should not be relied upon as the sole indicator of whether a child has a reading disorder. Fortunately, the shortcomings of phonemic awareness as a screening measure for long-term reading success are counterbalanced by another strength of phonemic awareness: it is highly amenable to change (Ehri et al., 2001; Suggate, 2016). For this reason, there is no reason that explicit instruction should not be included in intervention procedures for those children who demonstrate low phonemic awareness upon school entry. Because of the costs associated with providing intensive instruction to students who do not need it, however, schools should consider providing this instruction in a variety of ways depending on the needs of individual children.

Intervention

Previous meta-analyses conducted by Ehri et al. (2001) and Suggate (2016) have demonstrated that phonemic awareness instruction is effective for most children and that the effects can be sustained over time. For instance, Ehri et al. found that phonemic awareness instruction targeting segmentation demonstrated a strong effect on outcomes (d = .87) and instruction targeting blending resulted in a moderate effect on outcomes (d = .61). From this same review, Ehri and colleagues found that small group instruction was actually associated with a larger effect than individual instruction, which was good news for schools who generally lack the resources to provide individual instruction to large numbers of students. In a broad update of this literature in 2016, Suggate reported that phonemic awareness interventions were associated with small effects that were maintained over time. It should be noted, though, that his review was focused solely on studies that included follow up testing data and likely excluded the bulk of studies investigating phonemic awareness instruction since 2001.

For those children entering school who demonstrate poor phonological awareness and/or print knowledge, additional instructional support is likely necessary to catch them up to their peers who have had previous exposure to these concepts (Tiruchittampalam et al., 2018). Although supplemental instruction is likely necessary for these children, the intensity and amount of that instruction required for them to become successful readers will vary. This variability in necessary instructional support is critical to differentiating between students from disadvantaged backgrounds and those with reading disabilities. Response to Intervention (RTI) has been characterized as a tiered system of support for exactly these students. While newer iterations of RTI are commonly being referred to as multi-tiered systems of support (MTSS; e.g., (Eagle et al., 2015), the term RTI will be used in the present manuscript for conceptual systematicity and consistency. RTI as it is commonly implemented consists of tiered instruction designed to provide additional support to students who need it in the least restrictive environment possible. Tier one consists of high-quality instruction provided in the general education classroom to which most students should respond. Tier two, however, is often poorly defined in research and can take a variety of forms (e.g., Sterett et al., 2020), including small group instruction provided inside or outside of the general education classroom by a variety of personnel such as teachers, paraprofessionals, speech-language pathologists, or volunteers. In contrast, tier three instruction often consists of individual instruction for those students who did not respond adequately to instruction provided at the previous two levels. In some cases, tier three may be conceptualized as special education based on a Full and

Individual Evaluation of the child's needs. While RTI certainly has its drawbacks, including its inappropriate use to delay or prevent referrals to special education for children who likely have a disability (e.g., Office of Special Education and Rehabilitative Services, 2011), it still holds considerable value in providing high quality instruction to children while reserving intensive resource allocation to those who truly need them.

RTI is a preventative model of disability identification, developed in large part due to the inadequacies of other systems used to identify children with reading disorders (Justice, 2006). By providing additional instructional support before children are diagnosed with a reading disorder, RTI is an attempt to address the shortcomings of "wait to fail" models of disability identification such as those requiring a discrepancy between intellectual function and academic performance (e.g., Kranzler et al., 2019). Speech-language pathologists, with their specialized knowledge of language development (both spoken and written), are uniquely qualified to assist in the development and implementation of RTI services for children at risk of reading failure in the public schools (Justice, 2006; Ukrainetz, 2015). For the first tier of instruction, speech-language pathologists can provide instructional coaching to general education teachers and staff about typical language development as well as providing examples of efficient instructional practices in developing spoken and written language, especially in at-risk children. For those children who fail to adequately respond to tier one instruction, speech-language pathologists might help to design interventions that more specifically address their continued needs or provide direct services on a short-term basis via small group instruction. For those children who again fail to respond at this level or whose deficits are more apparently severe, a special education referral may be warranted or speech-language pathologists might consider more intensive instruction, again on a short-term basis supported by ongoing data collection. By collaborating with general education teachers and support staff throughout a school's RTI system, speech-language pathologists can leverage the depth of their knowledge in language development, assessment, and intervention practices to assist in the prevention of reading failure on their campuses.

CHAPTER THREE

Phonemic Awareness Instruction with At-Risk Readers: A Meta-Analysis

Phonemic awareness interventions have been previously demonstrated to be effective with many children but their effectiveness with children specifically identified as being at risk of reading failure remains unclear (Ehri et al., 2001; Suggate, 2016). Although previous meta-analyses have established that phonemic awareness instruction helps most children, they have also made it clear that this instruction is not equally effective for all children or across all outcome measures. For example, Ehri et al. found that instruction was not equally effective across measures of segmentation (d = .87), blending (d = .61), and deletion (d = .82) or the number of skills taught. They also reported no significant difference in intervention outcomes between children from lower socioeconomic backgrounds (d = 1.07) and their wealthier peers (d = 1.02) despite previous research suggesting that children from lower socioeconomic backgrounds enter school at a disadvantage relative to their wealthier peers (Bus et al., 1995).

Because phonemic awareness is critical to the acquisition of word decoding skills (Ehri, 2002), understanding how children already at risk of reading failure respond to instruction in this area is critical to attempts at preventing such failure. Effective early intervention is critical in preventing reading aversion, especially for children from lower socioeconomic backgrounds (Bus et al., 1995) and those from families with a history of language and literacy disorders (Catts et al., 2002) who are at greater risk for demonstrating reading problems. Children who become averse to reading experience fewer opportunities to engage successfully with the written word and the effects of these missed opportunities accumulate over time (Stanovich, 1986). Previous research suggests that children who are at risk of reading failure need instruction that is both more explicit and more intense than their peers in order to acquire the skills necessary for functional reading (Torgesen et al., 2003). This previous research has also demonstrated that interventions should specifically target the development of two phonemic awareness skills in particular to facilitate accurate word reading: segmentation and blending (Ehri et al., 2001). At the level of the individual phoneme, these two skills help children "sound out" unknown written words as well as spoken words whose written form is not yet familiar, resulting in a direct link to word decoding and encoding for children in the early stages of word recognition (Ehri, 2002).

Until children develop other, more efficient, strategies to word decoding such as recognizing words as a whole unit, comparing unknown words to known words, or recognizing familiar spelling patterns, phonemic decoding is necessary for accessing the written word (Kamhi & Catts, 2012). Facilitating the accuracy and fluency of such phonemic decoding skills is critical in reducing the cognitive load of word decoding by automatizing sound-symbol correspondences to some degree, enabling the development of more efficient strategies for word recognition, and thus allowing children to shift their focus to comprehension of what is being read (Perfetti, 1985).

The purpose of the present investigation is to update the previous work conducted by Ehri et al. (2001) and Suggate (2016) by clarifying the effects of phonemic awareness instruction with children at risk of reading failure. Specifically, this meta-analysis aims to quantify the effects of such interventions with regard to both theoretically relevant variables such as child age and socioeconomic background as well as practically relevant variables such as the incorporation of graphemes and the overall length of intervention. A better understanding of phonemic awareness instruction's effects with children at risk of reading failure is necessary in order to better facilitate appropriate curriculum development (Bus et al., 1995; Ehri et al., 2001; Suggate, 2016) and help assessment personnel differentiate between lack of educational opportunity and reading disabilities as required by the Individuals with Disabilities Education Improvement Act of 2004 (IDEA, 2004).

Method

Eligibility Criteria

In order to be considered for inclusion in this meta-analytic review, studies needed to report data from phonemic awareness interventions provided to children at risk of reading failure, especially those with previously documented decoding or comprehension scores below the 50th percentile, eligible for free or reduced school lunch programs, and/or having a documented family history of language and literacy disorders. Similar operationalization of risk was used in a previous review conducted by Ehri et al. (2001) and consistency with previous reviews allows consumers to better understand the development of the literature over time.

For inclusion in the review, studies must have reported (a) original data (b) collected using an experimental or quasi-experimental group research design (c) investigating the effects of phonemic awareness instruction (d) with children aged 18 years or younger who had intact sensory abilities and were (e) struggling or failing to read. Reports of these efforts must have been published in a peer-reviewed journal or available as a committee-approved dissertation. Studies must have also reported relevant effect sizes or the summary statistics required to calculate one (i.e., means, standard deviations, sample sizes). Finally, studies were only considered for inclusion if they were written in English.

Information Sources

The principal investigator and his research team searched the following databases to identify relevant studies: Academic Search Complete, ASHA Wire, Education Research Complete, ERIC, Medline, Open Dissertations, and Web of Science. Although there is substantial overlap in the journals indexed in these databases, this was desirable in order to maximize sensitivity at the cost of specificity (Card, 2012). The reference lists of included articles were also searched by the principal investigator for other relevant studies not included in the above databases and Google Scholar was used to seek out the works of authors whose work was included three or more times based on the database searches. In addition to the database searches, these backward and forward search attempts were critical in attempting to maximize the number of eligible studies identified for inclusion and reporting within this review. Additionally, the inclusion of data reported in committee-approved dissertations was intended to serve as some degree of protection against publication bias while also ensuring that data collection and reporting procedures were of sufficient quality to have been approved by the researcher's committee members.

Search Strategy

Database searches occurred between August and November of 2020 with forward and backward searches occurring in December 2020. The following terms were used for each database search, again attempting to maximize sensitivity at the cost of specificity as suggested by Card (2012): phonological awareness, phonemic awareness, intervention, instruction, at-risk, disadvantaged, learning difficulties, reading difficulties, RTI, and/or MTSS. These terms were curated from previous reviews (Ehri et al., 2001; Hall & Burns, 2018; Melby-Lervåg et al., 2012; Wren et al., 2013) again to help consumers understand how the literature has evolved over time. All databases were searched for the entirety of their coverage period.

Databases were searched by the principal investigator and two advanced graduate students with previous experience conducting systematic literature reviews. An initial search of these databases resulted in 2,696 articles being flagged for further review based on information contained in their title and abstract, with 1,643 unique reports identified. Of these 1,643 unique reports, 344 met the criteria for full-text review.

Selection Process

The titles and abstracts of the 1,643 unique reports identified through the database searches were each individually reviewed by one of the three research team members. Of these reports, 1,026 did not report the use of a phonological awareness intervention, 199 did not utilize a group research design, 26 reported the results of qualitative research, 33 were not journal articles or dissertations (e.g., book chapters), and 15 did not include children younger than 19 years of age. Journal articles and dissertations that suggested the investigation of a phonological awareness intervention implemented with children using a group research design were advanced for full-text review. Interrater reliability data were then collected by randomly selecting 333 of the 1,643 unique reports (20%) for a second round of review by the principal investigator. Interrater reliability was calculated by dividing the total number of rating agreements by the total number of opportunities for agreement. At the title-abstract review stage, interrater reliability was 96%. Disagreements were resolved through consensus.

The full texts of 344 journal articles and dissertations were then reviewed to determine which studies met inclusion criteria for this review. In order to be included in the meta-analysis, reports must have presented original data from a study using an experimental or quasi-experimental research design to investigate the use of a phonemic awareness intervention with children younger than 19 years of age demonstrating intact sensory abilities. Additionally, reports must have reported phonemic awareness outcomes in enough detail to readily calculate effect sizes or have reported an effect size that could be readily converted to Hedge's *g*. Of the 344 studies considered for inclusion, 82 did not report phonemic awareness outcomes, 81 didn't utilize a phonemic awareness intervention, 27 did not report or provide sufficient information to calculate effect sizes, 10 did not implement the intervention with children at-risk of reading failure, four did not include a control or experimental comparison group, three didn't report original data, three did not report data at all, and one included children with sensory impairments. After reviewing the full texts of these 344 reports, 133 met the present review's eligibility criteria and were advanced for data extraction. Interrater reliability data were then collected by randomly selecting 69 of the 344 reports (20%) for a second round of review. Interrater reliability was calculated in the same manner as with the title-abstract review stage. Interrater reliability at the full-text stage was calculated to be 91%. Disagreements were resolved through consensus.

Once a list of eligible reports obtained through database searches was complete, the principal investigator hand searched their reference lists to identify additional reports for consideration. These backward search procedures resulted in the identification of five additional reports meeting inclusion criteria for the present review; forward search procedures using Google Scholar resulted in the inclusion of no new reports. As a result of database searches, backward search procedures, and forward search procedures, a total of 138 studies were collected for review in the present meta-analysis.

Data Collection and Extraction

The principal investigator manually extracted data from the 138 included reports. If reports included multiple experiments, the one featuring the largest sample size was selected for inclusion. Once all data were extracted, the principal investigator then randomly selected 20% of the included studies and reviewed their extracted data to ensure the accuracy of the data extraction and entry processes; errors were found and corrected on two of the 28 studies selected for review in this manner.

Data were extracted from each included study based on relevant theory as well as the works of Ehri et al. (2001) and Troia (1999). First, reports were coded as to whether they represented a journal article or a dissertation. The year of publication or successful defense was also extracted. Interventions were coded as being either phonemic or phonological in scope, where phonological interventions represented those where some intervention time was allocated to work at less complex units such as the syllable or word levels. Additionally, interventions were coded as to whether they exclusively addressed phonemic or phonological awareness or also included procedures to address other literacy skills. As such, interventions were coded in one of four possible ways with regard to the complexity of their phonological work and incorporation of other literacy skills. The use of graphemes in the intervention was also recorded as a separate variable.

For each study, the mean age of all included participants was extracted or calculated based on the mean ages of each reported group; the mean ages and standard deviations for each individual group (i.e., treatment, control) were also separately extracted. The total sample size (*N*) for each study was extracted as well as the sample sizes of each reported group, with up to three groups reported per study. Data were extracted by group according to whether the group represented the study's experimental condition, control condition, or a secondary experimental condition. Specifically, studies that compared a standard phonemic awareness intervention to another intervention (e.g., phonemic awareness plus phonics instruction, phonological awareness instruction) were coded as having experimental versus secondary experimental conditions so as to differentiate them from studies comparing a phonemic awareness intervention to a no-treatment, irrelevant treatment (e.g., math interventions), or business-as-usual controls. Mean treatment group sizes and standard deviations were also extracted for all conditions. Means and standard deviations of group outcomes were also extracted for phonological awareness composites, segmentation, blending, deletion, and first phoneme identification.

Binary coding was used for the extraction of coding related to the data reported by each study. Specifically, each study was coded as to whether they reported an overall phonological awareness composite, reading composite, and/or outcomes related to spelling. Studies were also coded as to whether they reported each of the above-mentioned phonemic awareness skills (i.e., segmentation, blending, deletion, and first phoneme identification [FSID]), with an "other" variable added to account for studies that reported on additional phonological or phonemic awareness skills. Studies were then coded as to whether they reported distinct measures of word reading, nonword reading, and/or reading comprehension.

Studies' reporting of post-testing was also coded and up to two follow-up timepoints were noted, if applicable. Post-testing was coded as to whether it occurred immediately, between two and six months after the completion of the intervention, between seven and 26 months after the intervention, or at multiple points. If follow-up testing occurred, up to two of these timepoints were recorded as the number of months post intervention. Samples were coded as to the type of risk carried by the intervention group as well as the type of comparison group used. Possible risk factors included low socioeconomic status, screening measures (including teacher referral), speech or language impairments, or family history of language and literacy impairments as well as options for multiple risk factors or an otherwise unspecified risk category. Possible comparison group categories included typically developing peers of the same age, younger typically developing peers, similarly at-risk children to the relevant intervention group, children with previously identified disabilities, or "multiple" to represent the use of multiple comparison groups.

The grade level of at-risk children in each sample was extracted, with groupings for each grade pre-kindergarten through second or third grade and above. For studies reporting interventions with multiple grade levels, a fifth “multiple” option was used. The overall socioeconomic status of each sample was extracted as well, with options for low, middle/high, or mixed used for coding. The language of intervention was coded as either being provided in English or another language and the number of phonemic awareness skills taught during the intervention was reported as one, two, or three or more. The service model used to deliver each study’s intervention to participants was coded as individual, small group (two to five children), large group or classroom-based, or some combination of the above. For each study, the instructor was classified as being a speech-language pathologist, teacher, computer, parent, peer, researcher, other (often community-based volunteers or paraprofessionals), or some combination of the above. For coding purposes, anyone affiliated with the research team was coded as a researcher while the other categories were used to represent the credentials/status of anyone recruited to deliver the intervention. For example, if a member of the research team was a certified teacher and delivered the intervention, they were coded as a researcher. If, however, a classroom teacher was recruited to implement the intervention and was not reported to be a member of the research team, they were coded as a classroom teacher.

Dosage data were intended for extraction but were unavailable for almost all studies reviewed. In lieu of mean trial data for participants, the number of minutes per session were extracted as an approximation. As such, the number of minutes per session, the frequency per week of these sessions, and the duration of the intervention in weeks were extracted from each study. When studies reported the total number of sessions provided, this number was extracted as well. Cumulative intervention intensity was then calculated for each study by multiplying the number of individual session minutes, the number of sessions per week, and the total number of weeks covered by the intervention. Alternatively, if studies reported the total number of sessions participants received, cumulative intervention intensity was calculated by multiplying this number by the number of minutes included in each session.

Study Risk of Bias Assessment

For each included study, a quality assessment was conducted by coding methodological factors related to internal and external validity using the methods and definitions provided by Troia (1999). For a complete review of these variables and their definitions, the interested reader is referred to his original work. Troia's methods not only include these variables related to methodological rigor, but also a weighting system to reflect each variable's "importance to the causal interpretations and generalizability of a study" (p. 32). This weighting system was used to estimate the overall quality of each study, with summary statistics reported in Table 3.1. A few adjustments to Troia's criteria were made for the purpose of the current review, including substituting the reporting of risk criteria for his reporting of disability criteria as well as separating his participant selection variable into two variables: one representing the description of the population and another representing the explicit nature of inclusion and exclusion criteria. Both of these separated variables retained the same weight (three) as his original participant selection variable. Finally, Troia's original variable regarding expectations for transfer was omitted because this was not observed to be reported in the body of the manuscripts reviewed. Specifically, this variable was operationalized in his original work as "participants were informed that they would be expected to apply their learning in novel contexts" (p. 34). Its low weighting (one) in his original procedures also suggests that this was not intended to be a significant variable in interpreting the quality of reviewed studies.

Table 3.1

Study Quality Summary Statistics Using Troia's (1999) Weighted Scores

Category	Min	Med	*M*	*SD*	Max	Possible
Internal validity	3	20	19.92	5.74	34	36
General design	0	7	7.43	3.91	16	16
Measurement	0	8	7.50	2.84	11	11
Statistical treatment	0	5	4.99	1.21	7	9
External Validity	0	24	23.97	6.02	38	38
Research hypotheses	0	3	2.94	.44	3	3
Participant selection and recruitment	0	16	16.03	5.51	28	28
Transfer and maintenance measures	0	5	5.01	1.40	7	7
Overall Study Quality	3	45	43.89	9.29	66	74

Effect Measures

For each of the included studies, the means and standard deviations of groups' performances on those measures of phonemic awareness described above were extracted: overall ability, segmentation, blending, deletion, and FSID. These means, standard deviations, and group sizes were then used to hand calculate Hedge's *g* in Microsoft Excel using the formula: $\frac{M_1 - M_2}{SD_{pooled}}$. Metrics of a standardized mean difference between groups (e.g., Cohen's *d*, Hedge's *g*) are preferred in situations such as the present review when studies attempt to measure the same or similar outcomes using different instruments (Card, 2012). Hedge's *g* was selected for use in the present study due to its theoretical superiority to Cohen's *d* when working with small or significantly different sample sizes. As with Cohen's *d*, however, Hedge's *g* values of .20, .50, and .80 represent approximate benchmarks for small, medium, and large effects, respectively (Cohen, 1988).

A random sample of hand-calculated effect sizes was then compared to the results of an online calculator to ensure the accuracy of the calculations. For each of the above phonemic awareness outcomes, two potential effect sizes were calculated per report: one representing the difference between standard phonemic awareness interventions and interventions utilizing phonemic awareness instruction in some augmented or nonstandard way as well as one representing the difference between phonemic awareness instruction and no instruction. These hand-calculated results were then used to check the output of code run in the *R* statistical software throughout the analysis and synthesis portion of the meta-analytic review.

Synthesis Methods

In contrast to fixed-effect meta-analyses, random-effects meta-analyses allow for inferences to be made beyond the exact studies included in the review by assuming that the included studies are a representative sample of a much larger population of potential studies investigating a phenomenon (Card, 2012). In this way, the present review enables consumers to make inferences about the use of phonemic awareness interventions with at-risk children in general—not just the children included in the reviewed studies.

Previous reviews have investigated information regarding the effects of phonemic awareness instruction by aggregating outcomes within studies and then synthesizing those aggregates (Ehri et al., 2001). Although this practice protects the analysis against a violation of the assumption that each study contributes a single effect to the review, attempts at doing so in the present review resulted in significantly higher heterogeneity than when investigating outcomes separately. When aggregating the studies in the manner used similar to that used by Ehri et al. (2001), heterogeneity was considerable ($\tau^2 = .132$, $I^2 = 90.11\%$, $Q = 39.55$, $p < .001$) and valuable information about the impact of phonemic awareness interventions on specific skills was lost due to the aggregation. Attempts at pursuing subgroup analyses using aggregated data consistently resulted in nonsignificant results that were difficult to interpret in light of relevant theories. As such, aggregation of outcomes within single reports was not pursued in the present analysis and the effects of this decision are discussed later in the manuscript.

Studies were grouped according to a variety of theoretically and methodologically relevant characteristics in order to synthesize their outcomes. Because studies were included that compared two experimental conditions as well as those that compared an experimental condition to a control condition, this was used as primary a method of grouping studies for comparison. Several additional variables were also used for grouping studies for synthesis, including publication format, intervention type, overall sample size, study quality, mean age of participants, participant grade levels, socioeconomic status of the sample, instructor qualifications, delivery method, cumulative intervention intensity, the number of skills taught during the intervention, and the inclusion of graphemes during instruction.

Most reports reviewed and accepted for inclusion in the present review did not report gain scores and many reported summary statistics only for their post-testing outcomes. As such, a binary variable was created for each of the 139 included studies representing whether a study's groups started out equivalent at the onset of intervention. This variable was created to investigate the significance of groups' beginning interventions equivalent to each other as well as to assist with the creation of an interaction variable if these pre-intervention differences were significant.

The meta (Balduzzi et al., 2019) and metafor (Viechtbauer, 2010) packages as well as code reported by Schwarzer et al. (2015) were used to conduct a random-effects meta-analysis of the included studies using inverse variance weighting. As mentioned above, a random-effects model was pursued in order to make inferences beyond the specific studies included in the present review and based on previous research indicating significant variability in the effects of phonemic awareness interventions when used with children (Ehri et al., 2001).

Statistical heterogeneity was visualized using forest plots as well as quantified using multiple accepted measures, including Q, I^2, and τ^2 (Schwarzer et al., 2015). Cochrane's *Q* statistic represents the weighted sum of squared differences between each study's mean and the estimated effect, increasing with the number of included studies and affected by their sample sizes. In contrast, I^2 represents the percentage of variability in study estimates that is related to differences between the included studies rather than sampling error and is unaffected by the number of studies included in the analysis. However, it is also affected by the sample sizes of included studies in a similar fashion as Cochrane's *Q* statistic. Finally, τ^2 represents the underlying variability between studies and is systematically affected by neither the number of studies included nor their individual sample sizes. Because each of these metrics represents statistical heterogeneity in a slightly different manner, all three are reported for comparison in the present review.

The presence and significance of statistical heterogeneity was determined by interpreting the significance of Cochrane's *Q* statistic and the magnitude of I^2. Cochrane's guidelines for interpreting I^2 were used, with values between 0 and 40% representing heterogeneity unlikely to be significant, values between 30 and 60% representing moderate heterogeneity, values greater than or equal to 50% representing substantial heterogeneity, and values greater than or equal to 75% representing considerable heterogeneity between studies (Cochrane, 2021).

When significant statistical heterogeneity between studies was present, it was investigated through the use of subgroup analyses in order to clarify whether phonemic awareness outcomes varied systematically based on different study and sample characteristics (Card, 2012). These subgroup analyses were conducted based on the grouping variables discussed above (e.g., publication format, mean participant age, incorporation of graphemes). The significance of differences between groups was formally determined based on τ^2 values for each group in conjunction with the significance of Cochrane's *Q* statistic.

Results

Study Selection

A search of seven databases resulted in the identification of 2,696 reports, 1,643 of which were determined to be unique. Of these 1,643 unique reports, 334 were considered for inclusion and narrowed down to 133 that met inclusion for the present review. An additional five studies were identified through backward searches of the included studies' reference lists while no additional studies were found using forward search procedures. The flow of studies through the identification, screening, and inclusion stages of the review is visually summarized in Figure 3.1. Because specific data were not collected on the number of studies identified and screened using backward and forward search procedures, the final tally of studies included in the review (138) represents the final number of studies identified through database search procedures (133) plus those found through backward search procedures (5).

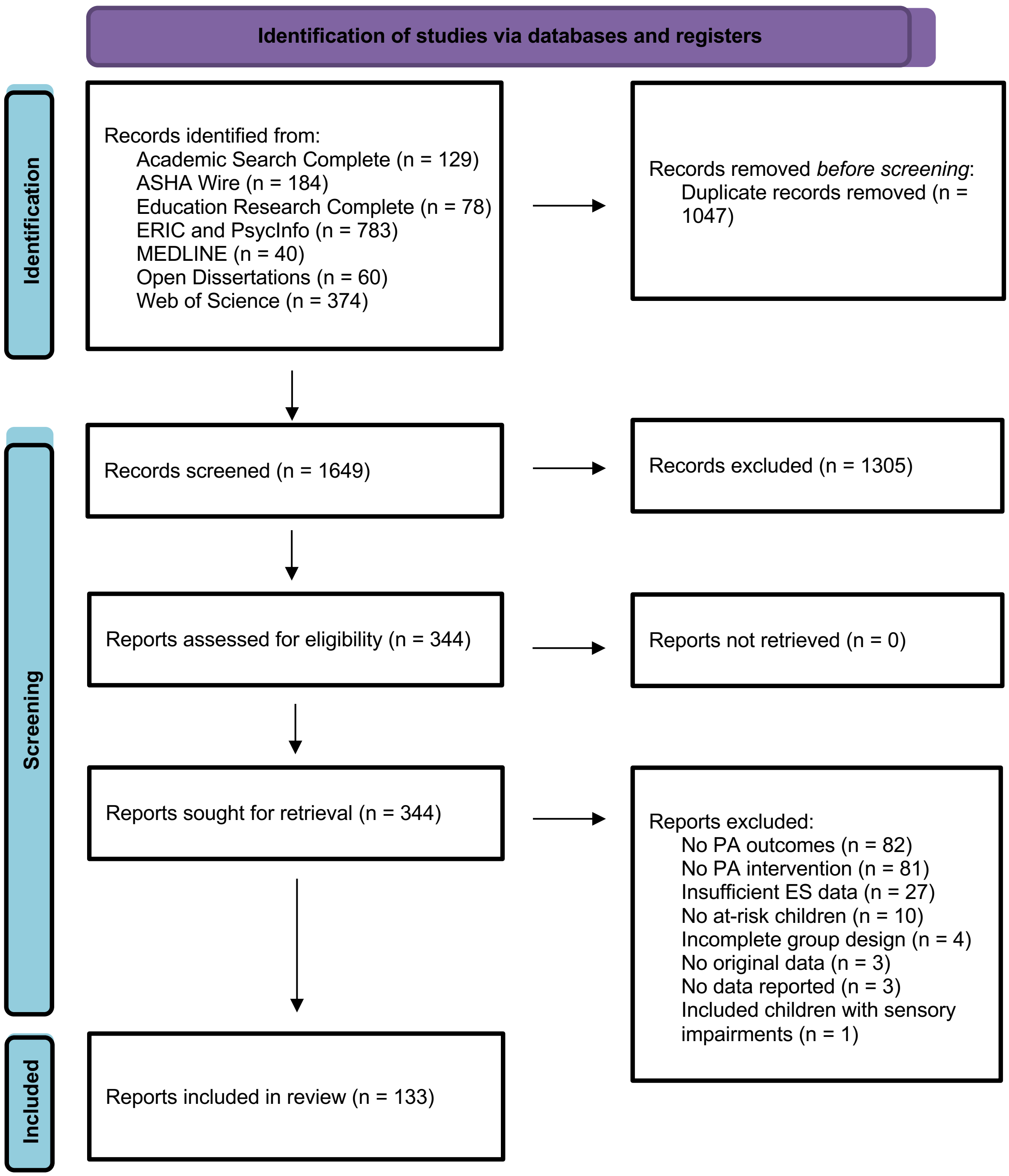

Figure 3.1. PRISMA flow diagram.

Study Characteristics

A complete list of studies included in this review is provided in Table 3.2 with citations provided in the Appendix. As is suggested in Table 3.2, many studies did not include enough information for all relevant variables to be calculated, including participants' mean age. The mean age of participants reported in this review is 6.36 years with a standard deviation of 1.72 across 115 articles and 23 dissertations. Forty studies investigated an intervention that spanned the phonological and phonemic levels, 53 spanned both levels of instruction and incorporated additional intervention components, 19 worked exclusively at the phonemic level, and 23 provided instruction at the phonemic level while also incorporating additional intervention components. Groups receiving no treatment (mostly characterized by continued access to general education instruction) were used in 78 studies, groups receiving an alternative treatment were used as a comparison in 20 studies, while 40 studies reported the use of both as comparison groups. Cumulative intervention intensity ranged from just under two hours (Chera & Wood, 2003) to approximately 117 hours (Hempenstall, 2008) with a mean of 26.52 and standard deviation of 24.22. Individual session length ranged from 10 minutes (Chera & Wood, 2003) to 100 minutes (Troia & Whitney, 2003) with a mean of 31.19 and standard deviation of 18.10. Sessions were conducted between one (Ukrainetz et al., 2009) and 10 times per week (Given et al., 2008) with a mean of 3.68 and 1.40. Graphemes were incorporated into the interventions of 118 studies, although it is noted that some studies did not report enough information to determine whether graphemes were included in their intervention procedures.

Table 3.2

Studies Included for Review

Manuscript			Intervention				
Author	Year	ID	Format	*N*	Age (*M*)	Type	Comparison
Adnams	2007	857	A	36	9.58	Phono+	Control
Al Otaiba	2005	1584	A	73	5.60	Phono+	Both
Allor	2004	876	A	157	6.57	Phono+	Control
Anderson	2010	982	D	61	4.50	Phono+	Alternative
Balbi	2020	574	A	125	7.04	Phono+	Control
Barker	1993	149	D	54	7.33	Phoneme	Both
Baugh	2005	1392	D	100		Phoneme+	Alternative
Benner	2003	179	D	36	5.70	Phono+	Control
Biancone	2018	1415	D	102	4.40	Phono	Control
Bingham	2010	1331	A	63		Phono	Control
Bjorn	2013	103	A	24		Phono	Control
Bode	2011	1037	A	307	5.58	Phoneme	Control
Bowyer-Crane	2008	774	A	137	4.75	Phono	Alternative
Brady	1994	1574	A	42	5.33	Phoneme+	Control
Calhoon	2007	524	A	76	6.54	Phono+	Control
Carson	2019b	260	A	23	5.43	Phono+	Control
Carson	2019	515	A	24	4.75	Phono+	Control
Chera	2003	188	A	75	4.08	Phoneme	Control
Compton	2000	943	A	41	6.56	Phono+	Control
Crespo	2018	396	A	530	6.31	Phono	Control
DeBaryshe	2007	164	A	81	3.93	Phono+	Control
Duff	2012	618	A	59	6.17	Phono	Control
Ehri	2007	1205	A	126	6.23	Phono	Alternative
Elbro	2004	909	A	82	6.17	Phoneme+	Control
Epstein	1994	105	D	167	4.42	Phono+	Control
Falth	2017	1055	A	27	6.00	Phono+	Control
Fives	2013	61	A	229	6.58	Phono+	Control
Flis	2018	102	D	82	6.38	Phono	Control
Foorman	1998	1523	A	133		Phono+	Both
Foy	2009	142	A	76		Phono+	Control
Fricke	2013	578	A	164	4.00	Phono+	Control
Fukuda	2013	628	D	66		Phoneme+	Control
Gillam	2008	1444	A	162	7.50	Phoneme+	Both
Gillon	2000	682	A	61	6.12	Phono	Both

(continued)

Manuscript					Intervention		
Author	Year	ID	Format	*N*	Age (*M*)	Type	Comparison
Gillon	2007	1056	A	20	6.39	Phono+	Alternative
Gillon	2005	1445	A	22	6.00	Phoneme+	Control
Given	2008	77	A	39	12.53	Phono+	Both
Gonzalez	2002	1227	A	52	9.79	Phono+	Both
Hadley	2000	284	A	22	5.75	Phono	Control
Hagans	2013	361	A	50		Phono+	Control
Hatcher	2004	131	A	304	4.67	Phono	Both
Hatcher	1994	579	A	93	7.49	Phoneme	Both
Hatcher	2006	666	A	77	5.50	Phono+	Alternative
Hecht	2002	590	A	76	5.58	Phono+	Control
Helf	2014	118	A	303		Phono+	Control
Hempenstall	2008	344	A	206	9.70	Phono+	Control
Hindson	2005	213	A	86		Phono	Control
Howell	2000	626	A	60		Phono	Control
Innes	2020	681	D	48	4.00	Phoneme	Both
Iverson	1993	1073	A	96	6.17	Phono+	Both
Jiménez	2003	409	A	59	8.74	Phoneme	Control
Kartal	2016	383	A	60	5.74	Phoneme+	Both
Kartal	2016	1570	A	20	5.21	Phono	Control
Kembert	2016	1569	A	424	4.57	Phono	Both
Kerins	2010	538	A	23	6.42	Phoneme+	Control
Korkman	1993	1667	A	46	6.12	Phoneme+	Control
Kyle	2013	212	A	30	5.83	Phoneme+	Both
Lane	2009	751	A	60		Phono+	Both
Lane	2007	980	A	24	6.59	Phoneme	Control
Layes	2020	310	A	44	10.27	Phono	Control
Layne	1997	740	D	16	7.62	Phoneme+	Control
Leafstedt	2004	513	A	62		Phoneme	Control
Lennox	2018	620	A	137	5.10	Phono	Control
Li	2018	1266	A	87	9.97	Phoneme	Control
Linan-Thompson	2005	1287	A	40	5.77	Phono	Control
Lo	2009	656	A	47		Phono+	Both
Loeb	2009	1421	A	74	7.42	Phoneme	Both
Lonigan	2016	31	A	132	4.85	Phono+	Control
Lonigan	2003	616	A	41	4.59	Phoneme	Control
Lonigan	2013	1240	A	203	4.53	Phono+	Both
MacKay	2011	211	D	100	5.35	Phono	Both

(continued)

Manuscript						Intervention	
Author	Year	ID	Format	N	Age (M)	Type	Comparison
Makhoul	2017	952	A	206		Phono	Alternative
Mathes	2001	1670	A	68	6.44	Phoneme+	Both
McMaster	2005	1237	A	56		Phono	Both
Milburn	2017	373	A	181	4.86	Phono+	Control
Mirbazel	2016	1390	A	60	12.50	Phoneme+	Control
Mitchell	2001	1417	A	72	6.34	Phoneme	Both
Nancollis	2005	1043	A	213	4.54	Phono	Control
Nelson	2005	176	A	36	5.70	Phono	Control
Nutkins	2004	649	D	55	4.63	Phoneme	Both
O'Connor	2010	30	A	69		Phono+	Control
O'Connor	2009	1021	A	54	6.67	Phono	Control
O'Shaugnessy	1997	1245	D	45	7.67	Phono+	Both
Olofsson	1985	631	A	83	6.92	Phoneme+	Control
Olson	1997	330	A	103	8.97	Phono+	Alternative
Pietrangelo	1999	976	D	129		Phono	Control
Pindiprolu	2020	1146	A	20		Phono	Alternative
Plasencia-Peinado	1999	648	D	102		Phono	Both
Pokorni	2004	1047	A	38	8.72	Phoneme+	Alternative
Poskiparta	1999	1644	A	30	7.13	Phoneme	Control
Pullen	2014	1338	A	100		Phoneme+	Both
Raisor	2006	20	D	44	3.50	Phoneme	Both
Regtvoort	2007	459	A	57	5.81	Phono+	Control
Ritter	2013	63	A	64	7.77	Phono	Control
Rule	2006	725	A	34		Phono	Both
Ryder	2008	665	A	24	6.71	Phono	Control
Scanlon	2008	1216	A	269		Phono	Alternative
Schneider	2000	1575	A	217		Phono+	Both
Schneider	1999	1668	A	90			Control
Schwartz	2005	901	A	72	6.41	Phono	Control
Segers	2004	332	A	36	5.75	Phoneme+	Both
Senechal	2012	1524	A	38	5.25	Phoneme+	Alternative
Seward	2009	619	D	34	5.75	Phono	Control
Simmons	2007	233	A	66	5.58	Phono+	Alternative
Solari	2018	435	A	90		Phono	Control
Solari	2008	651	A	34	5.50	Phono	Both
Stevens	2008	1179	A	462		Phono	Control
Torgesen	1999	1669	A	90	5.46		Alternative

(continued)

Manuscript					Intervention		
Author	Year	ID	Format	*N*	Age (*M*)	Type	Comparison
Torgeson	1997	325	A	105	5.43	Phono+	Both
Torgeson	2010	1147	A	108		Phono+	Both
Trainin	2014	1333	A	53		Phono+	Alternative
Travis	1997	548	D	64	6.50	Phoneme	Both
Troia	2003	13	A	37	9.39	Phono	Control
Ukrainetz	2000	174	A	36	5.48	Phoneme+	Control
Ukrainetz	2009	181	A	41	5.63	Phono+	Both
Vadasy	2006	290	A	67		Phono+	Control
Vadasy	1997	1396	A	40	6.54	Phono	Control
Vadasy	2000	1671	A	46	6.61		Control
Van Kleeck	1998	88	A	16	4.04	Phoneme	Alternative
van Otterloo	2009	416	A	57	5.87	Phono+	Control
Vanderwood	2014	195	A	105		Phoneme+	Control
Vaughn	2006	509	A	89	6.60	Phono	Control
Vaughn	2006	511	A	41	6.59	Phono+	Control
Wang	2000	880	D	47	7.50	Phono	Both
Warren	2009	1439	D	10	5.50	Phoneme	Control
Warrick	1993	999	A	28		Phono+	Control
Weiner	1994	555	A	40		Phoneme+	Both
Wheldall	2016	575	A	240	5.42	Phono+	Control
Whittaker	2013	1241	D	264		Phono+	Both
Wise	2015	81	A	87	6.25	Phoneme	Both
Wise	2016	789	D	12	5.75	Phono+	Control
Wise	1999	1054	A	110	8.98	Phoneme+	Both
Wise	2000	1573	A	200	8.94	Phono+	Alternative
Wolff	2011	528	A	112	9.25	Phono+	Control
Wolff	2014	1172	A	112	9.25	Phono+	Control
Yeh	2008	152	A	84	4.75	Phono+	Alternative
Yeh	2003	559	A	44	5.08	Phoneme+	Alternative
Zoski	2015	841	D	12	6.01	Phono+	Alternative

Risk of Bias in Studies

Study quality varied substantially and a summary of the internal validity, external validity, and overall quality of each study based on procedures developed by Troia (1999) are reported in Table 3.3. Studies were rank ordered based on their overall weighted quality score that was calculated as a combination of their internal and external validity scores. For a review of the weighting procedures, interested readers are referred to Table 1 in Troia (1999). A coding error occurred when recording whether studies utilized the unit of treatment as the unit of their statistical analysis and this variable was subsequently excluded from all analyses in the present manuscript. Furthermore, these study quality scores should be interpreted with some caution as much of the scoring is based on whether studies report information rather than the quality of the information reported. For example, these procedures only required studies to report that “a procedure was used to ensure treatment conditions were being implemented faithfully” (p. 33). This reporting varied from researchers mentioning that implementers were observed (Wheldall et al., 2016) to weekly observation and recording of implementation using an intervention components checklist with data then reported in the manuscript (O'Shaugnessy, 1997). As such, information being reported did not always correlate with the quality of the procedures themselves. Such variability in coding procedures is likely to have increased the heterogeneity present in the data, which is discussed below.

Table 3.3

Summary of Overall Study Quality

Manuscript		Weighted Scores			
Author	Year	Internal Validity	External Validity	Overall Quality	Rank
O'Shaugnessy	1997	31	35	66	1
Torgesen	1999	31	33	64	2
Gillam	2008	32	30	62	3
Loeb	2009	25	36	61	4
Ritter	2013	28	30	58	5
Lane	2007	26	32	58	5
Korkman	1993	20	38	58	5
Troia	2003	21	36	57	6
Wise	2016	23	34	57	6
Gillon	2000	21	35	56	7
Bowyer-Crane	2008	25	31	56	7
Pullen	2014	28	28	56	7
Ukrainetz	2009	29	25	54	8
Kyle	2013	28	26	54	8
Pokorni	2004	22	32	54	8
McMaster	2005	29	25	54	8
Al Otaiba	2005	21	33	54	8
Mathes	2001	24	30	54	8
Hatcher	2004	25	28	53	9
Olson	1997	27	26	53	9
Calhoon	2007	23	30	53	9
Hatcher	1994	20	33	53	9
Gillon	2007	26	27	53	9
Torgeson	2010	26	27	53	9
Trainin	2014	34	19	53	9
Biancone	2018	23	29	52	10
Senechal	2012	28	24	52	10
Fives	2013	26	25	51	11
Barker	1993	22	29	51	11
Simmons	2007	32	19	51	11
Jiménez	2003	20	31	51	11
Solari	2018	24	27	51	11
Li	2018	23	28	51	11
Nelson	2005	24	26	50	12
Segers	2004	26	24	50	12
Regtvoort	2007	16	34	50	12
Zoski	2015	30	20	50	12
O'Connor	2010	27	22	49	13
Yeh	2008	30	19	49	13
Nutkins	2004	22	27	49	13
Lane	2009	27	22	49	13
Wolff	2014	17	32	49	13
Ehri	2007	19	30	49	13

(continued)

Manuscript		Weighted Scores			
Author	Year	Internal Validity	External Validity	Overall Quality	Rank
Vadasy	2000	22	27	49	13
Given	2008	25	23	48	14
Bjorn	2013	21	27	48	14
Vadasy	2006	22	26	48	14
Layes	2020	23	25	48	14
Hatcher	2006	23	25	48	14
Lonigan	2013	26	22	48	14
Mitchell	2001	22	26	48	14
Lonigan	2016	23	24	47	15
Epstein	1994	18	29	47	15
van Otterloo	2009	16	31	47	15
Yeh	2003	28	19	47	15
Balbi	2020	18	29	47	15
Fricke	2013	20	27	47	15
Solari	2008	29	18	47	15
Schwartz	2005	15	32	47	15
Wise	1999	22	25	47	15
Wise	2000	23	24	47	15
MacKay	2011	22	24	46	16
Vaughn	2006	22	24	46	16
Travis	1997	25	21	46	16
Plasencia-Peinado	1999	18	28	46	16
Gillon	2005	22	24	46	16
Foorman	1998	21	25	46	16
Benner	2003	21	24	45	17
Lonigan	2003	16	29	45	17
Ryder	2008	16	29	45	17
Adnams	2007	16	29	45	17
Allor	2004	18	27	45	17
Scanlon	2008	17	28	45	17
Whittaker	2013	16	29	45	17
Warren	2009	24	21	45	17
Wang	2000	20	24	44	18
Pindiprolu	2020	29	15	44	18
Carson	2019b	16	27	43	19
Hagans	2013	20	23	43	19
Weiner	1994	21	22	43	19
Fukuda	2013	19	24	43	19
Elbro	2004	12	31	43	19
Wise	2015	20	22	42	20
Hadley	2000	15	27	42	20
Milburn	2017	18	24	42	20
Vaughn	2006	18	24	42	20
Wolff	2011	23	19	42	20
Lo	2009	17	25	42	20
Baugh	2005	23	19	42	20

(continued)

Manuscript		Weighted Scores			
Author	Year	Internal Validity	External Validity	Overall Quality	Rank
Mirbazel	2016	19	22	41	21
Vadasy	1997	17	24	41	21
Poskiparta	1999	15	26	41	21
DeBaryshe	2007	16	24	40	22
Ukrainetz	2000	20	20	40	22
Vanderwood	2014	16	24	40	22
Hindson	2005	15	25	40	22
Torgeson	1997	20	20	40	22
Carson	2019	16	24	40	22
Hecht	2002	18	22	40	22
Anderson	2010	21	19	40	22
Helf	2014	15	24	39	23
Kerins	2010	17	22	39	23
Gonzalez	2002	14	25	39	23
Leafstedt	2004	20	18	38	24
Bingham	2010	18	20	38	24
Van Kleeck	1998	15	22	37	25
Nancollis	2005	10	27	37	25
Pietrangelo	1999	17	19	36	26
Kembert	2016	15	21	36	26
Kartal	2016	18	18	36	26
Duff	2012	11	24	35	27
Seward	2009	13	22	35	27
Iverson	1993	13	22	35	27
Flis	2018	15	19	34	28
Foy	2009	15	19	34	28
Crespo	2018	15	19	34	28
Wheldall	2016	21	13	34	28
Lennox	2018	16	18	34	28
Layne	1997	17	17	34	28
Schneider	1999	10	24	34	28
Raisor	2006	14	19	33	29
Chera	2003	15	18	33	29
Rule	2006	22	11	33	29
Bode	2011	12	21	33	29
Falth	2017	11	22	33	29
Brady	1994	6	27	33	29
Linan-Thompson	2005	11	21	32	30
Hempenstall	2008	13	18	31	31
Compton	2000	11	20	31	31
Innes	2020	18	12	30	32
Stevens	2008	18	12	30	32
Howell	2000	16	13	29	33
Warrick	1993	15	14	29	33
Schneider	2000	16	13	29	33
Kartal	2016	17	11	28	34

(continued)

Manuscript		Weighted Scores			
Author	Year	Internal Validity	External Validity	Overall Quality	Rank
Makhoul	2017	12	15	27	35
Olofsson	1985	12	12	24	36
O'Connor	2009	3	0	3	38

Results of Individual Studies

Outliers for each outcome were identified using the boxplot function in *R*, which operationalizes outliers as data points outside the boundaries of ± 1.5 times the interquartile range of the dataset. These outliers were checked to ensure their accuracy as an outlier rather than a product of a data entry error. Relevant outliers were removed from the dataset prior to data synthesis (discussed below) and the boxplots (including outliers) for all five outcomes are presented in Figure 3.2.

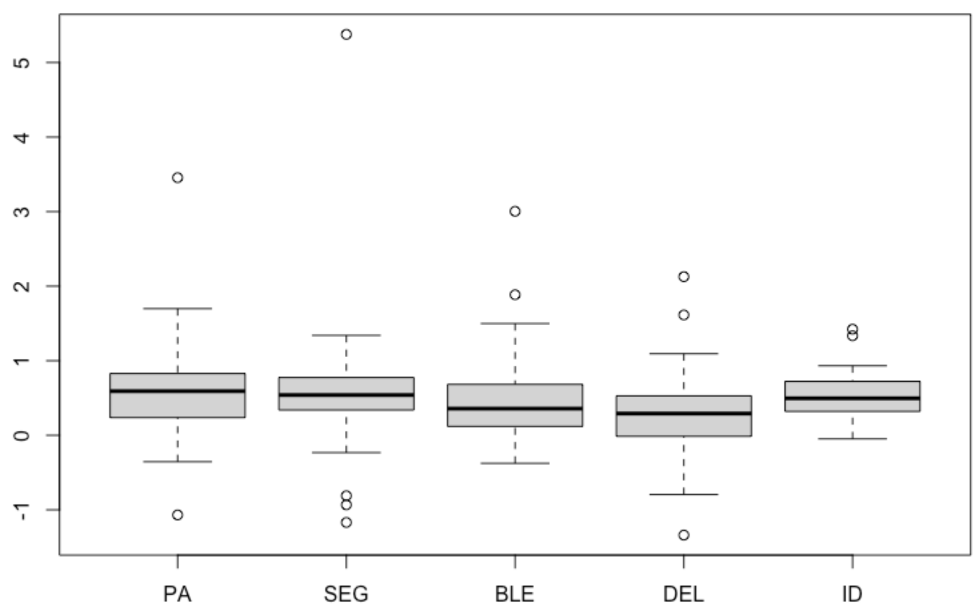

Figure 3.2. Boxplots of phonemic awareness outcomes.

Results of Syntheses

Forest plots for all measured study outcomes (phonological awareness, segmentation, blending, deletion, and FSID) after the removal of outliers are presented in Figures 3.3-3.7 and include each study's experimental and control sample sizes, means, and standard deviations. Additionally, these forest plots allow for easy visualization and comparison of each study's effect size and associated 95% confidence interval. Studies were synthesized and heterogeneity within the dataset was investigated through the use of subgroup analyses. Subgroups were investigated both through visual comparison of their effects as well as through formal statistical tests of the heterogeneity between subgroups. These subgroup analyses should be interpreted with caution, however, as there were frequently fewer than 10 studies per subgroup (Cochrane, 2021). Regardless, the estimated effect sizes for each subgroup provide some insight into the impact of different variables on the outcomes of phonemic awareness interventions.

Study	Experimental Total	Experimental Mean	Experimental SD	Control Total	Control Mean	Control SD	SMD	95%-CI	Weight (fixed)	Weight (random)
Balbi 2020	68	15.24	6.5900	57	12.35	6.1900	0.45	[0.09; 0.80]	3.4%	2.9%
Benner 2003	18	100.60	9.4000	18	90.90	9.6000	1.00	[0.30; 1.70]	0.9%	1.8%
Compton 2000	21	10.95	2.0100	20	11.50	0.8200	-0.35	[-0.97; 0.27]	1.1%	2.0%
DeBaryshe 2007	51	13.20	7.0000	30	8.50	6.4000	0.69	[0.22; 1.15]	2.0%	2.5%
Duff 2012	29	9.83	4.2500	30	11.17	4.0900	-0.32	[-0.83; 0.20]	1.6%	2.3%
Foorman 1998	41	2.16	0.8300	53	1.53	0.8800	0.73	[0.31; 1.15]	2.4%	2.7%
Fricke 2013	82	8.42	4.1100	82	7.55	4.3200	0.21	[-0.10; 0.51]	4.6%	3.1%
Gillon 2000	23	46.29	13.0100	15	29.41	19.1900	1.05	[0.36; 1.75]	0.9%	1.8%
Given 2008	12	3.42	11.5200	13	5.69	10.8600	-0.20	[-0.98; 0.59]	0.7%	1.5%
Gonzalez 2002	17	56.41	3.4600	17	45.37	12.8700	1.14	[0.41; 1.88]	0.8%	1.7%
Hempenstall 2008	134	17.04	3.4700	72	15.28	4.3600	0.46	[0.17; 0.75]	5.1%	3.2%
Innes 2020	16	42.30	9.4000	16	36.40	11.0000	0.56	[-0.15; 1.27]	0.9%	1.7%
Kartal 2016	22	25.27	11.8500	22	15.15	12.6000	0.81	[0.20; 1.43]	1.1%	2.0%
Kartal 2016	10	42.40	13.4000	10	35.50	14.5600	0.47	[-0.42; 1.36]	0.5%	1.3%
Kembert 2016	109	70.06	18.0300	187	57.19	19.0500	0.69	[0.44; 0.93]	7.3%	3.3%
Korkman 1993	26	0.50	0.7000	20	0.20	0.8000	0.40	[-0.19; 0.98]	1.2%	2.1%
Lane 2009	17	53.41	3.6900	24	49.38	7.2000	0.66	[0.02; 1.30]	1.1%	1.9%
Lane 2007	13	84.63	12.3300	11	76.40	8.5000	0.74	[-0.10; 1.57]	0.6%	1.4%
Layes 2020	22	10.59	5.6800	22	5.68	4.0700	0.98	[0.35; 1.60]	1.1%	2.0%
Layne 1997	9	16.80	2.5000	7	15.20	1.2000	0.74	[-0.29; 1.77]	0.4%	1.1%
Leafstedt 2004	16	7.63	6.5500	46	4.69	3.7200	0.63	[0.05; 1.21]	1.3%	2.1%
Lonigan 2013	67	31.16	14.2200	72	29.00	13.0800	0.16	[-0.18; 0.49]	3.9%	3.0%
Milburn 2017	91	94.38	10.3900	90	91.22	9.1700	0.32	[0.03; 0.61]	5.0%	3.2%
Mirbazel 2016	30	20.13	0.9100	30	17.26	2.6700	1.42	[0.85; 1.99]	1.3%	2.1%
Mitchell 2001	24	72.50	10.3100	24	63.10	16.0300	0.69	[0.10; 1.27]	1.3%	2.1%
Nelson 2005	18	9.23	6.4000	18	1.89	5.7600	1.18	[0.46; 1.89]	0.8%	1.7%
O'Connor 2009	28	9.11	2.8600	26	8.62	3.2800	0.16	[-0.38; 0.69]	1.5%	2.3%
Poskiparta 1999	15	24.90	6.7000	15	23.50	9.2000	0.17	[-0.55; 0.89]	0.8%	1.7%
Raisor 2006	17	25.70	11.7000	10	13.90	8.5000	1.07	[0.23; 1.91]	0.6%	1.4%
Regtvoort 2007	31	35.50	14.3000	26	26.10	17.6000	0.58	[0.05; 1.12]	1.5%	2.3%
Rule 2006	10	18.90	6.0000	14	8.20	6.5000	1.64	[0.68; 2.60]	0.5%	1.2%
Schneider 1999	57	22.81	9.7600	33	18.58	10.7800	0.41	[-0.02; 0.85]	2.3%	2.6%
Seward 2009	17	31.25	7.5100	17	21.06	9.0900	1.19	[0.46; 1.93]	0.8%	1.7%
Solari 2008	20	53.80	14.9600	7	51.83	11.8000	0.13	[-0.73; 1.00]	0.6%	1.4%
Solari 2018	55	16.89	3.5000	35	15.57	6.3700	0.27	[-0.15; 0.70]	2.4%	2.7%
Stevens 2008	247	87.37	47.9000	215	73.67	44.8000	0.29	[0.11; 0.48]	12.7%	3.5%
Ukrainetz 2009	14	30.50	6.4000	13	25.20	7.5000	0.74	[-0.04; 1.52]	0.7%	1.5%
Vadasy 2006	36	88.00	11.9000	31	85.00	10.2000	0.27	[-0.22; 0.75]	1.8%	2.5%
van Otterloo 2009	30	50.72	24.4000	27	43.61	28.8500	0.26	[-0.26; 0.79]	1.6%	2.3%
Vaughn 2006	22	58.32	15.2000	19	45.75	17.4000	0.76	[0.12; 1.40]	1.1%	1.9%
Vaughn 2006	42	69.33	14.3000	47	55.48	18.6000	0.82	[0.39; 1.26]	2.3%	2.6%
Wheldall 2016	115	32.93	8.6900	125	31.26	13.1300	0.15	[-0.11; 0.40]	6.7%	3.3%
Whittaker 2013	153	8.92	1.5200	86	9.33	1.1900	-0.29	[-0.56; -0.02]	6.1%	3.3%
Wise 1999	42	10.00	3.5000	31	5.10	3.0000	1.47	[0.95; 1.99]	1.6%	2.3%
Wolff 2014	57	10.20	4.6600	55	10.20	4.0400	0.00	[-0.37; 0.37]	3.1%	2.9%
Fixed effect model	**1994**			**1838**			**0.41**	**[0.34; 0.47]**	**100.0%**	**--**
Random effects model							**0.51**	**[0.38; 0.64]**	**--**	**100.0%**

Heterogeneity: $I^2 = 68\%$, $\tau^2 = 0.1095$, $p < 0.01$

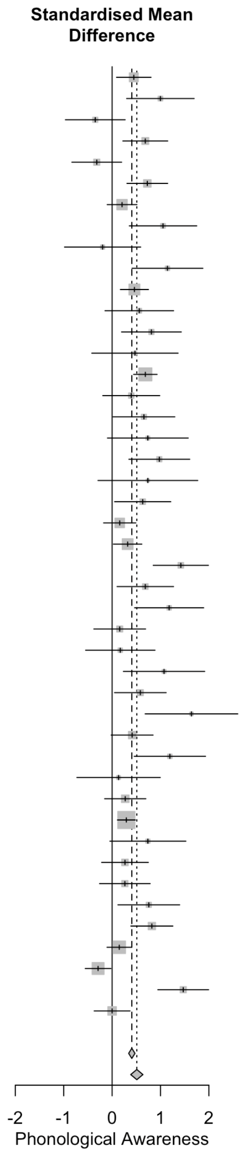

Figure 3.3. Forest plot for studies reporting phonological awareness composite scores.

A total of 45 studies investigated the effect of phonemic awareness instruction on phonological and/or phonemic awareness composite outcomes between children that received instruction and those who did not. These 45 studies reported outcomes for 1,994 children who received intervention and 1,838 who received no or business-as-usual instruction. Even after the removal of two outliers (Howell et al., 2000; Ryder et al., 2008), there was still substantial heterogeneity present between the remaining studies (τ^2

= .110, I^2 = 67.9%, Q = 137.22, $p < .001$) with phonemic awareness instruction resulting in a medium effect on composite outcomes (g = .511, $p < .001$). Subgroup analyses were pursued to investigate this heterogeneity beginning with the equivalency of comparison groups pre-intervention and these analyses are summarized in Table 3.4. Statistically significant differences between subgroups were observed at the .05 level for the type of intervention provided (Q = 10.1, p = .018) as well as the use of graphemes in the intervention (Q = 4.84, p = .028). Interventions including instruction at both the phonological and phonemic levels resulted in a medium effect on composite outcomes both with (g = .593) and without (g = .624) other literacy skills being addressed. The effect of phonemic awareness instruction alone resulted in a large effect (g = 1.002) while phonemic awareness instruction incorporating the teaching of other literacy skills resulted in only a small effect (g = .340). Interventions including graphemes resulted in a medium effect on composite outcomes (g = .545) while those not using graphemes had only a small effect (g = .250). There was not a statistically significant difference across outcomes for studies based on their cumulative intervention intensity. A summary of study outcomes grouped by their percentile rank regarding cumulative intervention intensity is included in Table 3.5.

Table 3.4

Subgroup Analyses of Phonemic Awareness Composite Outcomes

Variable	Group	*n*	g	*Q*	*P*
Starting equivalency				1.84	.175
	No	15	.366		
	Yes	28	.565		
Manuscript type				0.26	.609
	Dissertation	6	.672		
	Article	39	.498		
Sample size				0.26	.612
	< 30	3	.647		
	≥ 30	42	.507		
Publication date (2001)				0.32	.571
	Before 2001	9	.594		
	After 2001	36	.487		
Socioeconomic status				0.32	.852
	Low	19	.434		
	Mid/High	6	.521		
	Mixed	8	.512		
Participant age (years)				4.26	.749
	4	4	.528		
	5	5	.346		
	6	12	.614		
	7	7	.402		
	8	1	.740		
	9	2	.723		
	10	3	.765		
	13	2	.636		
Participant grade				1.51	.912
	Preschool	9	.419		
	Kindergarten	14	.496		
	First	9	.522		
	Second	1	.739		
	Third and up	3	.261		
	Mixed	7	.579		
Intervention type				10.1	.018
	Phono	6	.624		
	Phono+	13	.593		
	Phone	5	1.002		
	Phone+	20	.340		

(continued)

Variable	Group	*n*	g	*Q*	*P*
Intervention provider				4.79	.571
	SLP	2	.914		
	Teacher	14	.441		
	Computer	4	.499		
	Parent	2	.420		
	Peer	0	—		
	Researcher	14	.551		
	Other	5	.304		
	Combination	2	.942		
Delivery model				3.13	.372
	Individual	12	.530		
	Small group	16	.468		
	Large group	8	.635		
	Combination	4	.239		
Number of skills taught				4.45	.108
	1	1	.728		
	2	7	.346		
	3 or more	30	.551		
Graphemes included				4.84	.028
	No	2	.250		
	Yes	40	.545		

Table 3.5

Outcomes by Cumulative Intervention Intensity Percentile Rank

Outcome	<25%	25% < x < 50%	50% < x < 75%	75% < x < 100%	*Q*	*p*
Composite	.721	.330	.431	.418	4.97	.174
Segmentation	.525	.612	.423	.661	3.03	.388
Blending	.402	.152	.230	.306	1.82	.612
Deletion	.179	.136	.425	.305	2.19	.534
FSID	.389	.783	.462	.449	2.01	.570

A total of 43 studies reported enough information to determine whether their groups started out equivalent when measuring outcomes using a composite score. Starting equivalency alone did not explain the heterogeneity between studies and there was not a significant difference between the 15 studies whose groups did not start out equivalent (g = .366) and the 28 studies whose groups did (g = .565) where Q = 1.84 and p = .175.

Starting equivalency was planned to be used as an interaction variable for the remainder of the subgroup analyses but its lack of significance suggested that this would not contribute to meaningful results and would also likely obscure the remaining analyses by reducing the size of already small groups.

Prior to the removal of outliers, there was considerable heterogeneity within the 44 studies reporting segmentation outcomes when comparing children who received phonemic awareness instruction to those who did not (τ^2 = .254, I^2 = 80.3%, Q = 218.77, p < .001). After removal of these four outliers (Lo et al., 2009; Nancollis et al., 2005; Ryder et al., 2008; Vanderwood et al., 2014), there was only minimal heterogeneity remaining in the data (τ^2 = .019, I^2 = 23.3%, Q = 50.82, p = .097). A forest plot of the remaining 40 studies reporting segmentation outcomes is presented in Figure 3.4. A random-effects meta-analysis of these 40 studies indicated an average Hedge's *g* of .571, representing a medium effect of phonemic awareness instruction on segmentation outcomes for children at-risk of reading failure. Although there was not significant heterogeneity in the remaining data, subgroup data are reported for descriptive purposes in Table 3.6.

Study	Experimental Total	Experimental Mean	Experimental SD	Control Total	Control Mean	Control SD	SMD	95%-CI	Weight (fixed)	Weight (random)
Adnams 2007	18	65.00	32.9000	18	68.33	28.4900	-0.11	[-0.76; 0.55]	1.4%	1.8%
Allor 2004	76	30.52	17.1000	81	23.46	14.7600	0.44	[0.12; 0.76]	5.8%	5.1%
Barker 1993	17	7.56	3.7000	18	4.61	3.9000	0.76	[0.07; 1.45]	1.2%	1.6%
Benner 2003	18	17.10	8.1000	18	11.20	14.6000	0.49	[-0.18; 1.15]	1.3%	1.7%
Brady 1994	21	0.76	1.5500	21	0.13	0.4400	0.54	[-0.07; 1.16]	1.5%	1.9%
Calhoon 2007	43	22.87	16.3700	33	14.51	14.2500	0.53	[0.07; 1.00]	2.7%	3.1%
Carson 2019	14	5.86	2.3100	10	3.21	2.2900	1.11	[0.23; 1.99]	0.7%	1.0%
Crespo 2018	321	24.06	11.6000	209	14.95	12.0100	0.77	[0.59; 0.95]	17.8%	8.4%
Foy 2009	37	32.72	15.1800	39	23.81	14.9600	0.59	[0.13; 1.05]	2.7%	3.1%
Gillon 2005	12	1.70	2.5000	10	0.90	1.5000	0.36	[-0.48; 1.21]	0.8%	1.1%
Hagans 2013	26	53.88	8.7000	24	39.33	17.6500	1.04	[0.45; 1.64]	1.6%	2.1%
Hatcher 1994	30	18.63	6.0900	31	14.94	7.6600	0.53	[0.01; 1.04]	2.2%	2.6%
Hecht 2002	42	7.91	7.0500	34	1.53	2.8400	1.13	[0.64; 1.62]	2.4%	2.8%
Helf 2014	111	50.87	18.6000	192	43.49	17.2300	0.41	[0.18; 0.65]	10.4%	6.9%
Iverson 1993	32	16.88	4.5300	32	12.69	6.8200	0.71	[0.21; 1.22]	2.3%	2.7%
Jiménez 2003	31	9.42	0.9200	28	8.42	2.0200	0.64	[0.12; 1.16]	2.1%	2.5%
Kartal 2016	22	60.00	31.6200	22	50.91	28.2700	0.30	[-0.30; 0.89]	1.6%	2.1%
Kartal 2016	10	0.50	0.7100	10	0.20	0.4100	0.50	[-0.40; 1.39]	0.7%	1.0%
Kerins 2010	12	8.91	1.9200	11	6.56	4.4500	0.67	[-0.17; 1.52]	0.8%	1.1%
MacKay 2011	25	31.92	16.9600	46	13.93	14.7300	1.14	[0.62; 1.67]	2.1%	2.5%
Mathes 2001	19	4.95	4.6000	24	5.42	4.7600	-0.10	[-0.70; 0.50]	1.6%	2.0%
McMaster 2005	21	21.90	12.4000	20	20.20	13.0300	0.13	[-0.48; 0.74]	1.5%	2.0%
Mitchell 2001	24	18.70	3.2000	24	16.60	3.2100	0.64	[0.06; 1.23]	1.7%	2.1%
Nelson 2005	18	12.61	7.2800	18	4.38	10.6100	0.88	[0.20; 1.57]	1.2%	1.6%
O'Connor 2010	38	38.59	17.7700	31	27.43	17.0200	0.63	[0.15; 1.12]	2.4%	2.9%
Olofsson 1985	51	15.30	2.9000	32	14.30	4.3000	0.28	[-0.16; 0.73]	2.9%	3.3%
Plasencia-Peinado 1999	36	27.10	12.2000	34	22.80	13.2000	0.33	[-0.14; 0.81]	2.6%	3.0%
Schneider 2000	48	3.59	2.6200	115	1.98	2.6200	0.61	[0.27; 0.96]	4.9%	4.6%
Schwartz 2005	36	17.70	4.9300	36	15.27	5.4300	0.46	[0.00; 0.93]	2.6%	3.0%
Segers 2004	12	0.90	1.8000	12	0.00	0.0000			0.0%	0.0%
Torgeson 1997	33	9.20	3.8000	36	7.10	3.2000	0.59	[0.11; 1.08]	2.5%	2.9%
Torgeson 2010	35	15.60	3.7000	39	11.70	4.5000	0.93	[0.45; 1.41]	2.5%	2.9%
Travis 1997	21	10.76	3.1100	21	10.45	3.6200	0.09	[-0.52; 0.70]	1.6%	2.0%
Ukrainetz 2000	18	5.70	2.5000	18	2.00	3.0000	1.31	[0.58; 2.04]	1.1%	1.5%
Ukrainetz 2009	14	5.30	2.2000	13	3.60	1.8000	0.82	[0.03; 1.61]	0.9%	1.3%
Vadasy 1997	20	16.75	3.6700	20	14.65	6.0300	0.41	[-0.21; 1.04]	1.5%	1.9%
Vadasy 2000	23	15.51	3.7900	23	11.15	5.5300	0.90	[0.29; 1.51]	1.6%	2.0%
van Otterloo 2009	30	8.93	6.1700	27	7.24	6.7100	0.26	[-0.26; 0.78]	2.1%	2.6%
Warren 2009	5	35.00	29.7700	5	42.00	30.5000	-0.21	[-1.45; 1.04]	0.4%	0.5%
Warrick 1993	14	6.40	13.4000	14	2.10	5.8000	0.40	[-0.35; 1.15]	1.0%	1.4%
Weiner 1994	13	9.62	0.5100	13	9.31	1.1800	0.33	[-0.44; 1.11]	1.0%	1.3%
Fixed effect model	**1447**			**1462**			**0.58**	**[0.51; 0.66]**	**100.0%**	**--**
Random effects model							**0.57**	**[0.48; 0.67]**	**--**	**100.0%**

Heterogeneity: $I^2 = 23\%$, $\tau^2 = 0.0190$, $p = 0.10$

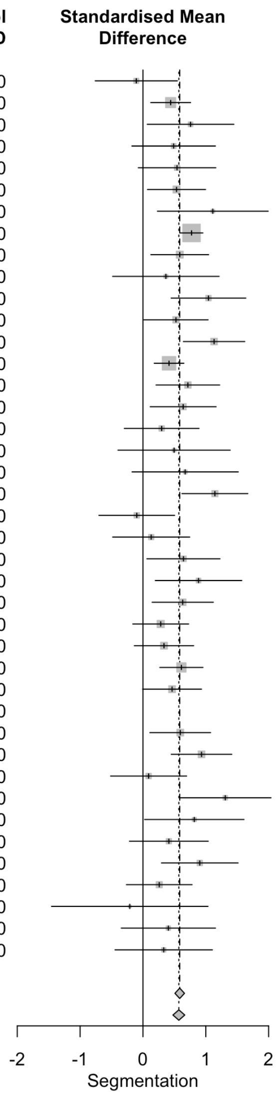

Figure 3.4. Forest plot for studies reporting segmentation scores.

Table 3.6

Subgroup Analyses of Segmentation Scores

Variable	Group	*n*	g	*Q*	*p*
Starting equivalency				.05	.827
	No	11	.535		
	Yes	24	.562		
Manuscript type				.12	.730
	Dissertation	6	.511		
	Article	34	.578		
Sample size				.06	.805
	< 30	6	.527		
	≥ 30	34	.574		
Publication date (2001)				.92	.337
	Before 2001	16	.513		
	After 2001	24	.605		
Socioeconomic status				.68	.711
	Low	14	.529		
	Mid/High	5	.578		
	Mixed	9	.635		
Participant age (years)				9.37	.052
	5	6	.849		
	6	14	.587		
	7	8	.466		
	9	1	.640		
	10	1	-.106		
Participant grade				10.37	.035
	Preschool	3	.953		
	Kindergarten	15	.587		
	First	16	.500		
	Second	0	—		
	Third and up	1	-.106		
	Mixed	5	.734		
Intervention type				1.89	.596
	Phono	6	.500		
	Phono+	8	.617		
	Phone	8	.421		
	Phone+	18	.603		

(continued)

Variable	Group	*n*	g	*Q*	*p*
Intervention provider				11.96	.102
	SLP	3	.320		
	Teacher	7	.614		
	Computer	8	.645		
	Parent	2	.189		
	Peer	3	.233		
	Researcher	7	.691		
	Other	4	.514		
	Combination	4	.856		
Delivery model				.69	.876
	Individual	18	.543		
	Small group	12	.605		
	Large group	6	.623		
	Combination	1	.365		
Number of skills taught				.16	.923
	1	3	.524		
	2	5	.626		
	3 or more	26	.545		
Graphemes included				.33	.568
	No	5	.511		
	Yes	33	.594		

Statistically significant differences were observed regarding the effects of intervention by participants' grade (Q = 10.37, p = .035), with younger children seeing the largest gains. Specifically, phonemic awareness instruction provided to preschool-aged children resulted in a large effect on segmentation outcomes (g = .953) while instruction provided to children across multiple grades exhibited the next largest, though medium, effect (g = .734). One study enrolling children in the third grade and up category demonstrated a negligible negative effect related to children who did not receive instruction (g = -.106). The Adnams et al. (2007) for which this negligible negative effect was observed did not demonstrate group starting equivalency, with control participants demonstrating a substantial advantage on the segmentation measure at the start of intervention.

There were 37 studies that reported phonemic blending outcomes as a method of comparing children who received phonemic awareness instruction to those who did not. Before the removal of outliers, there was substantial heterogeneity present between these studies ($\tau^2 = .107$, $I^2 = 62.2\%$, $Q = 95.11$, $p < .001$). After the removal of two outliers (Carson, 2020; Ryder et al., 2008), the heterogeneity was reduced to moderate levels but remained statistically significant ($\tau^2 = .066$, $I^2 = 51.3\%$, $Q = 69.75$, $p < .001$). A small effect of phonemic awareness instruction was observed with regard to blending outcomes ($g = .341$, $p < .001$). Subgroup analyses were pursued to investigate this heterogeneity and a summary of these analyses is presented in Table 3.7.

Study	Experimental Total	Experimental Mean	Experimental SD	Control Total	Control Mean	Control SD	SMD	95%-CI	Weight (fixed)	Weight (random)
Adnams 2007	18	79.06	26.8700	18	78.20	28.5600	0.03	[-0.62; 0.68]	1.6%	2.3%
Al Otaiba 2005	24	9.08	2.2200	24	7.79	2.1300	0.58	[0.00; 1.16]	2.0%	2.7%
Barker 1993	17	7.22	4.2000	18	7.22	4.4000	0.00	[-0.66; 0.66]	1.5%	2.3%
Bjorn 2013	13	-0.28	0.4200	11	-0.33	0.9900	0.07	[-0.74; 0.87]	1.0%	1.8%
Bode 2011	150	8.32	2.2600	157	6.69	2.4000	0.70	[0.47; 0.93]	12.6%	5.2%
Falth 2017	19	7.39	3.0600	8	6.00	2.3200	0.47	[-0.37; 1.31]	1.0%	1.7%
Flis 2018	41	13.88	2.4900	41	14.76	2.1800	-0.37	[-0.81; 0.06]	3.5%	3.6%
Gillam 2008	54	9.00	3.9000	54	7.90	3.6000	0.29	[-0.09; 0.67]	4.7%	4.0%
Hatcher 1994	30	14.43	5.1600	31	11.03	7.1100	0.54	[0.03; 1.05]	2.6%	3.1%
Hecht 2002	42	10.29	5.5500	34	4.24	5.0800	1.12	[0.63; 1.61]	2.8%	3.2%
Innes 2020	16	4.50	5.4000	16	4.50	4.9000	0.00	[-0.69; 0.69]	1.4%	2.2%
Kartal 2016	22	49.75	27.5600	22	32.90	21.4300	0.67	[0.06; 1.28]	1.8%	2.5%
Kartal 2016	10	0.50	0.7100	10	0.40	0.7000	0.14	[-0.74; 1.01]	0.9%	1.6%
Kerins 2010	12	9.27	1.1900	11	8.67	2.2900	0.32	[-0.50; 1.15]	1.0%	1.7%
Linan-Thompson 2005	21	1.76	1.6700	19	1.10	1.6000	0.40	[-0.23; 1.02]	1.7%	2.5%
Loeb 2009	24	7.42	3.4600	25	6.24	3.1400	0.35	[-0.21; 0.92]	2.1%	2.8%
Lonigan 2003	20	1.95	2.1300	21	2.50	2.1400	-0.25	[-0.87; 0.36]	1.8%	2.5%
Lonigan 2016	66	13.81	4.0500	66	13.26	4.1000	0.13	[-0.21; 0.48]	5.8%	4.3%
Lonigan 2013	67	14.33	7.1200	72	14.07	6.9100	0.04	[-0.30; 0.37]	6.1%	4.4%
McMaster 2005	21	9.71	6.7000	20	11.95	7.5600	-0.31	[-0.92; 0.31]	1.8%	2.5%
Milburn 2017	91	91.07	12.6700	90	89.55	12.8500	0.12	[-0.17; 0.41]	7.9%	4.7%
Mitchell 2001	24	15.20	3.6000	24	13.00	4.8000	0.51	[-0.07; 1.09]	2.0%	2.7%
Olofsson 1985	51	14.60	3.3000	32	11.60	4.0000	0.83	[0.37; 1.29]	3.2%	3.4%
Pietrangelo 1999	67	6.79	3.1500	62	5.23	2.4500	0.55	[0.19; 0.90]	5.4%	4.2%
Poskiparta 1999	15	6.20	3.8000	15	4.30	3.9000	0.48	[-0.25; 1.21]	1.3%	2.0%
Ritter 2013	34	9.88	2.0400	30	8.00	2.3100	0.86	[0.34; 1.37]	2.5%	3.1%
Schneider 2000	48	5.43	1.7700	115	4.80	1.9700	0.33	[-0.01; 0.67]	5.9%	4.3%
Segers 2004	12	2.30	3.2000	12	0.40	0.7000	0.79	[-0.04; 1.63]	1.0%	1.7%
Torgeson 1997	33	18.00	4.9000	36	16.30	5.4000	0.33	[-0.15; 0.80]	3.0%	3.3%
Torgeson 2010	35	20.60	4.5000	39	18.20	5.4000	0.48	[0.01; 0.94]	3.1%	3.4%
Travis 1997	21	18.90	4.6600	21	20.32	5.1900	-0.28	[-0.89; 0.33]	1.8%	2.6%
Troia 2003	25	9.72	2.7800	12	9.08	2.4300	0.23	[-0.46; 0.92]	1.4%	2.2%
Ukrainetz 2009	14	8.70	1.5000	13	6.40	3.4000	0.86	[0.07; 1.66]	1.1%	1.8%
van Otterloo 2009	30	10.50	5.8000	27	9.34	7.0200	0.18	[-0.34; 0.70]	2.5%	3.0%
Wise 2016	5	13.00	0.2610	7	8.50	3.8700	1.38	[0.05; 2.72]	0.4%	0.8%
Fixed effect model	**1192**			**1213**			**0.35**	**[0.27; 0.44]**	**100.0%**	**--**
Random effects model							**0.34**	**[0.22; 0.47]**	**--**	**100.0%**

Heterogeneity: $I^2 = 51\%$, $\tau^2 = 0.0658$, $p < 0.01$

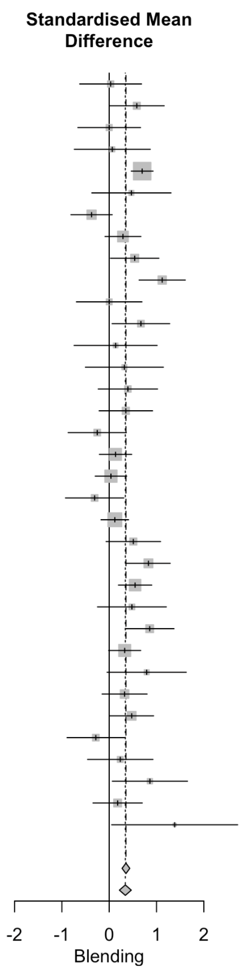

Figure 3.5. Forest plot for studies reporting blending scores.

Table 3.7

Subgroup Analyses of Blending Scores

Variable	Group	*n*	g	*Q*	*p*
Starting equivalency				.61	.436
	No	8	.221		
	Yes	23	.355		
Manuscript type				1.59	.207
	Dissertation	6	.100		
	Article	29	.388		
Sample size				0	.963
	< 30	5	.349		
	≥ 30	30	.339		
Publication date (2001)				.39	.531
	Before 2001	9	.403		
	After 2001	26	.323		
Socioeconomic status				.09	.957
	Low	15	.313		
	Mid/High	3	.258		
	Mixed	4	.263		
Participant age (years)				10.42	.108
	4	1	.000		
	5	6	.101		
	6	13	.531		
	7	6	.350		
	8	2	.546		
	9	1	.234		
	10	1	.030		
Participant grade				14.81	.005
	Preschool	7	.165		
	Kindergarten	11	.577		
	First	9	.164		
	Second	0	—		
	Third and up	2	.044		
	Mixed	3	.589		
Intervention type				3.3	.347
	Phono	9	.274		
	Phono+	8	.209		
	Phone	5	.541		
	Phone+	13	.357		

(continued)

Variable	Group	*n*	g	*Q*	*p*
Intervention provider				16.15	.013
	SLP	3	.585		
	Teacher	6	.556		
	Computer	12	.321		
	Parent	1	.179		
	Peer	1	-.308		
	Researcher	8	.131		
	Other	0	—		
	Combination	3	.391		
Delivery model				.13	.936
	Individual	15	.332		
	Small group	12	.287		
	Large group	6	.344		
	Combination	0	—		
Number of skills taught				.97	.323
	1	0	—		
	2	6	.184		
	3 or more	24	.327		
Graphemes included				.17	.677
	No	5	.258		
	Yes	29	.337		

Statistically significant differences between groups were observed by participant grade level (Q = 14.81, p = .005) and intervention provider (Q = 16.16, p = .013), again with several groups containing fewer than 10 studies. Kindergarten students and those enrolled across multiple grades demonstrated the greatest gains from the provision of phonemic awareness instruction with medium effects of .577 and .589 on average, respectively. Medium effects of phonemic awareness instruction on blending outcomes were also observed when intervention was provided by a speech-language pathologist (g = .585) or teacher (g = .556), with a single, small negative effect observed when instruction was provided by a peer (g = -.308). The McMaster et al. (2005) study from which this single negative effect was obtained did not demonstrate starting equivalency, with control participants demonstrating a practically significant advantage (*d* = .49) over their intervention counterparts at the beginning of the study.

Three outliers were observed in and removed from the data of studies reporting deletion scores after children were exposed to phonemic awareness instruction (O'Shaugnessy, 1997; Wise, 2016; Wise et al., 1999). There was moderate heterogeneity present in the remaining 31 studies ($\tau^2 = .053$, $I^2 = 46.5\%$, $Q = 56.05$, $p < .05$) that was then investigated through the use of subgroup analyses; a forest plot of these 31 studies is presented in Figure 3.6. A small effect on deletion outcomes was observed across these 31 studies comparing 1,066 children who received instruction compared to 1,147 who did not or received business-as-usual instruction ($g = .248$, $p < .001$). The results of these subgroup analyses are presented in Table 3.8; formal statistical analyses of group differences should be interpreted with some caution given the small sample sizes within each (Cochrane, 2021).

Study	Experimental Total	Experimental Mean	Experimental SD	Control Total	Control Mean	Control SD	SMD	95%-CI	Weight (fixed)	Weight (random)
Barker 1993	17	4.50	3.2000	18	2.89	3.0000	0.51	[-0.17; 1.18]	1.6%	2.5%
Bjorn 2013	13	0.27	1.5600	11	0.00	1.5500	0.17	[-0.64; 0.97]	1.1%	1.9%
Bode 2011	150	1.29	1.5000	157	1.20	1.5700	0.06	[-0.17; 0.28]	14.6%	6.4%
Brady 1994	21	3.46	2.9300	21	2.67	2.4200	0.29	[-0.32; 0.90]	2.0%	2.8%
Elbro 2004	35	3.60	4.2000	47	2.10	3.0000	0.42	[-0.03; 0.86]	3.7%	4.0%
Flis 2018	41	9.34	3.4500	41	10.27	4.3000	-0.24	[-0.67; 0.20]	3.9%	4.1%
Hadley 2000	11	6.18	2.8600	11	3.91	2.7000	0.79	[-0.09; 1.66]	1.0%	1.7%
Hatcher 2004	102	16.52	4.5200	89	14.38	6.7700	0.38	[0.09; 0.66]	8.9%	5.7%
Hatcher 1994	30	13.70	6.1700	31	7.39	6.2000	1.01	[0.47; 1.54]	2.6%	3.3%
Iverson 1993	32	11.00	5.8500	32	9.43	4.9800	0.29	[-0.21; 0.78]	3.0%	3.6%
Kartal 2016	22	34.09	36.3400	22	37.73	38.0400	-0.10	[-0.69; 0.50]	2.1%	2.9%
Kartal 2016	10	0.20	0.4200	10	0.10	0.3200	0.26	[-0.62; 1.14]	0.9%	1.6%
Kyle 2013	10	18.50	6.1000	10	12.20	5.4000	1.05	[0.10; 2.00]	0.8%	1.5%
Loeb 2009	24	4.62	3.5400	25	4.28	3.3500	0.10	[-0.46; 0.66]	2.3%	3.1%
Lonigan 2003	20	0.48	1.2500	21	2.10	2.5700	-0.78	[-1.42; -0.14]	1.8%	2.6%
Lonigan 2016	66	10.27	3.6800	66	8.60	3.7700	0.45	[0.10; 0.79]	6.1%	5.0%
Lonigan 2013	67	9.12	6.5700	72	7.93	6.1300	0.19	[-0.15; 0.52]	6.6%	5.1%
Mathes 2001	19	4.14	4.0600	24	4.18	2.0400	-0.01	[-0.61; 0.59]	2.0%	2.9%
Milburn 2017	91	97.68	12.0600	90	92.91	10.4800	0.42	[0.13; 0.71]	8.4%	5.6%
Poskiparta 1999	15	4.90	3.9000	15	5.40	4.6000	-0.11	[-0.83; 0.60]	1.4%	2.3%
Ryder 2008	12	8.25	1.7600	12	6.50	2.1100	0.87	[0.03; 1.71]	1.0%	1.8%
Schwartz 2005	36	6.64	2.5600	36	5.58	2.5000	0.41	[-0.05; 0.88]	3.4%	3.8%
Torgeson 1997	33	15.80	5.7000	36	13.00	4.9000	0.52	[0.04; 1.00]	3.2%	3.7%
Torgeson 2010	35	15.30	4.2000	39	12.50	4.6000	0.63	[0.16; 1.10]	3.3%	3.8%
Travis 1997	21	11.76	3.1600	21	13.18	4.8700	-0.34	[-0.95; 0.27]	2.0%	2.8%
Troia 2003	25	7.64	2.2300	12	7.67	1.7800	-0.01	[-0.70; 0.67]	1.5%	2.4%
Ukrainetz 2000	18	2.90	3.2000	18	1.60	2.7000	0.43	[-0.23; 1.09]	1.7%	2.5%
Wang 2000	15	19.80	0.4100	17	17.53	3.7600	0.80	[0.08; 1.53]	1.4%	2.2%
Weiner 1994	13	7.15	1.6300	13	7.00	2.2000	0.08	[-0.69; 0.84]	1.2%	2.0%
Wise 2015	5	10.00	4.8000	75	12.73	4.9700	-0.54	[-1.45; 0.36]	0.9%	1.6%
Wolff 2011	57	4.74	1.4700	55	4.82	1.4000	-0.06	[-0.43; 0.32]	5.3%	4.7%
Fixed effect model	**1066**			**1147**			**0.24**	**[0.16; 0.33]**	**100.0%**	**--**
Random effects model							**0.25**	**[0.12; 0.37]**	**--**	**100.0%**

Heterogeneity: $I^2 = 46\%$, $\tau^2 = 0.0528$, $p < 0.01$

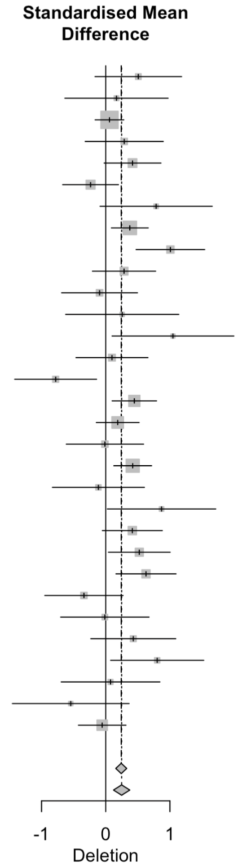

Figure 3.6. Forest plot for studies reporting deletion scores.

Table 3.8

Subgroup Analyses of Deletion Scores

Variable	Group	*n*	g	*Q*	*p*
Starting equivalency				.73	.393
	No	9	.352		
	Yes	19	.200		
Manuscript type				.23	.634
	Dissertation	4	.138		
	Article	27	.268		
Sample size				1.57	.211
	< 30	4	.512		
	≥ 30	27	.228		
Publication date (2001)				1.25	.264
	Before 2001	12	.354		
	After 2001	19	.197		
Socioeconomic status				.88	.644
	Low	12	.289		
	Mid/High	4	.013		
	Mixed	3	.318		
Participant age (years)				6.06	.194
	5	9	.292		
	6	10	.163		
	7	6	.331		
	8	1	.802		
	9	2	-.046		
Participant grade				56.05	.003
	Preschool	6	.248		
	Kindergarten	5	.216		
	First	11	.096		
	Second	2	.892		
	Third and up	2	-.016		
	Mixed	3	.496		
Intervention type				2.43	.488
	Phono	8	.022		
	Phono+	9	.316		
	Phone	7	.263		
	Phone+	7	.331		

(continued)

Variable	Group	*n*	g	*Q*	*p*
Intervention provider				10.28	.068
	SLP	0	—		
	Teacher	9	.337		
	Computer	9	.035		
	Parent	0	—		
	Peer	1	-.013		
	Researcher	7	.177		
	Other	1	.870		
	Combination	3	.603		
Delivery model				3.77	.287
	Individual	13	.194		
	Small group	12	.334		
	Large group	4	.101		
	Combination	2	.415		
Number of skills taught				3.21	.201
	1	1	1.047		
	2	4	.155		
	3 or more	21	.215		
Graphemes included				.08	.779
	No	7	.205		
	Yes	23	.253		

In reviewing these subgroup analyses, statistically significant differences were only observed by participant grade ($Q = 56.05$, $p = .003$) where second grade students' outcomes indicated a large effect of intervention ($g = .892$) while first and third grade or older students demonstrated negligible change compared to controls with *g* values of .096 and -.016, respectively. The grouped negative outcome demonstrated by third grade participants was calculated based on data contributed by Björn & Leppänen (2013) and Wolff (2011) with effects of .168 and -.055, respectively; Wolff's participants did not begin the study with equivalent levels of deletion performance.

Of the 23 studies reporting FSID outcomes, two were identified as outliers and removed from the dataset (Hindson et al., 2005; Nelson et al., 2005). Surprisingly, there was no heterogeneity observed between the remaining 21 studies ($\tau^2 = 0$, $I^2 = 0\%$, $Q =$

19.25, p = .506) and a small effect of phonemic awareness instruction was observed (g = .428, p < .001). There was not a significant difference between studies whose groups were equivalent pre-intervention (Q = 2.82, p = .093) and a summary of subgroup analyses is reported for descriptive purposes in Table 3.9 with a forest plot of all relevant studies presented in Figure 3.7.

Study	Experimental Total	Experimental Mean	Experimental SD	Control Total	Control Mean	Control SD	SMD	95%-CI	Weight (fixed)	Weight (random)
Adnams 2007	18	97.62	6.4300	18	93.06	23.5700	0.26	[-0.40; 0.91]	3.1%	3.1%
Benner 2003	18	16.40	7.9000	18	9.30	7.6000	0.90	[0.21; 1.58]	2.8%	2.8%
Biancone 2018	51	6.04	3.8800	51	4.82	3.9500	0.31	[-0.08; 0.70]	8.8%	8.8%
Bingham 2010	38	8.21	2.4300	25	7.34	2.7200	0.34	[-0.17; 0.85]	5.2%	5.2%
Carson 2019	14	5.71	2.8100	10	5.80	1.6000	-0.04	[-0.85; 0.78]	2.0%	2.0%
Chera 2003	60	9.60	3.8000	15	7.90	4.8000	0.42	[-0.15; 0.99]	4.1%	4.1%
Elbro 2004	35	8.60	3.7000	47	6.70	3.3000	0.54	[0.10; 0.99]	6.8%	6.8%
Fricke 2013	82	5.83	3.7000	82	5.46	3.5600	0.10	[-0.20; 0.41]	14.3%	14.3%
Fukuda 2013	24	1.90	1.6600	42	1.48	1.0000	0.33	[-0.18; 0.83]	5.3%	5.3%
Innes 2020	16	11.80	3.0000	16	10.00	3.1000	0.58	[-0.13; 1.28]	2.7%	2.7%
Kartal 2016	22	89.46	15.5800	22	78.18	28.0500	0.49	[-0.11; 1.09]	3.7%	3.7%
Kartal 2016	10	6.20	3.9900	10	6.40	4.2700	-0.05	[-0.92; 0.83]	1.7%	1.7%
Lennox 2018	110	5.40	2.4000	27	4.40	2.4000	0.41	[-0.01; 0.84]	7.5%	7.5%
Linan-Thompson 2005	21	3.48	0.9800	19	3.37	1.1200	0.10	[-0.52; 0.72]	3.5%	3.5%
MacKay 2011	25	31.52	15.4400	46	23.89	15.3700	0.49	[0.00; 0.98]	5.5%	5.5%
Nutkins 2004	23	15.39	9.9400	10	11.97	8.8300	0.35	[-0.40; 1.09]	2.4%	2.4%
Raisor 2006	17	4.50	4.5000	10	1.60	2.7000	0.71	[-0.10; 1.52]	2.1%	2.1%
Schneider 2000	48	6.13	2.2800	115	3.69	3.1400	0.83	[0.48; 1.18]	11.0%	11.0%
Ukrainetz 2000	18	9.00	1.3000	18	6.60	3.4000	0.91	[0.22; 1.60]	2.8%	2.8%
Ukrainetz 2009	14	9.90	0.5000	13	9.40	1.0000	0.62	[-0.16; 1.40]	2.2%	2.2%
Warrick 1993	14	17.10	21.3000	14	6.40	0.8400	0.69	[-0.08; 1.46]	2.3%	2.3%
Fixed effect model	**678**			**628**			**0.43**	**[0.31; 0.54]**	**100.0%**	**--**
Random effects model							**0.43**	**[0.31; 0.54]**	**--**	**100.0%**

Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.51$

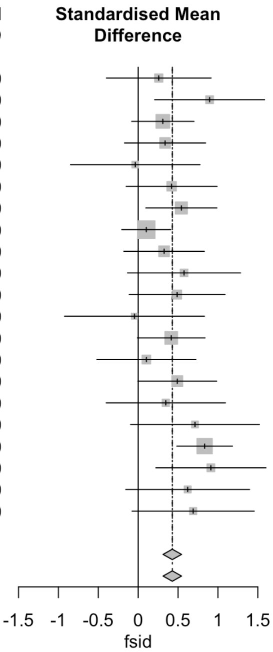

Figure 3.7. Forest plot for studies reporting FSID scores.

Table 3.9

Subgroup Analyses of FSID Scores

Variable	Group	*n*	g	*Q*	*p*
Starting equivalency				2.82	.093
	No	9	.535		
	Yes	11	.332		
Manuscript type				.08	.778
	Dissertation	7	.457		
	Article	14	.419		
Sample size				.66	.412
	< 30	3	.232		
	≥ 30	18	.440		
Publication date (2001)				8.74	.003
	Before 2001	18	.826		
	After 2001	3	.351		
Socioeconomic status				.22	.895
	Low	10	.465		
	Mid/High	1	.337		
	Mixed	5	.460		
Participant age (years)				2.09	.555
	4	5	.279		
	5	6	.411		
	6	5	.513		
	10	1	.258		
Participant grade				2.88	.578
	Preschool	7	.324		
	Kindergarten	10	.539		
	First	1	.488		
	Second	0	—		
	Third and up	1	.258		
	Mixed	2	.367		
Intervention type				1.67	.643
	Phono	4	.494		
	Phono+	6	.328		
	Phone	4	.526		
	Phone+	7	.478		

(continued)

Variable	Group	*n*	g	*Q*	*p*
Intervention provider				2.06	.841
	SLP	2	.409		
	Teacher	3	.558		
	Computer	4	.292		
	Researcher	6	.554		
	Other	3	.355		
	Combination	3	.391		
Delivery model				3.42	.332
	Individual	5	.408		
	Small group	5	.502		
	Large group	6	.544		
	Combination	2	.222		
Number of skills taught				6.16	.046
	1	1	.309		
	2	6	.201		
	3 or more	12	.534		
Graphemes included				0	.977
	No	2	.435		
	Yes	19	.429		

Statistically significant differences in FSID outcomes were observed only for study publication year ($Q = 8.74$, $p = .003$) and the number of skills taught during the intervention ($Q = 6.16$, $p = .046$). It should be noted that substantially more studies addressed FSID as an outcome prior to 2001, which may have influenced the results. Interventions incorporating three or more skills were associated with a medium effect ($g = .534$) while small effects were noted for interventions targeting only one ($g = .309$) or two ($g = .201$) skills.

Reporting Biases

In order to evaluate the relationship between study effects and their precision, two tests were conducted for each outcome using the metabias function in *R* (Begg & Mazumdar, 1994; Egger et al., 1997). A summary of these tests is presented in Table 3.10

and the evidence suggests that small-study effects and/or publication bias are unlikely to have affected outcomes except for the phonological awareness composite. For phonological awareness composite scores, both tests suggest that there is a relationship between a study's size and the observed effect associated with that study and that synthesis of these studies is likely to be biased (Schwarzer et al., 2015).

Table 3.10

Outcomes of Tests for Small Study Effects

	Begg & Mazumdar (1994)		Egger et al. (1997)	
Outcome	*z*	*p*	*t*	*p*
Composite	2.61	.009	2.94	.005
Segmentation	.07	.944	-.80	.426
Blending	.30	.766	-.28	.779
Deletion	.15	.878	.26	.799
FSID	1.33	.184	1.29	.214

Discussion

This meta-analysis investigated the effect of phonemic awareness instruction on multiple outcomes, including phonological awareness composites, segmentation, blending, deletion, and FSID scores. Previous reviews have indicated that phonemic awareness instruction can be effective for many children (Ehri et al., 2001; Suggate, 2016) but the present review specifically investigated their use with children at risk of reading failure.

Summary of the Evidence

Previous reviews have demonstrated that the effect of phonemic awareness instruction varies from small (d = .39; Suggate, 2016) to large (d = .86; Ehri et al., 2001) depending on a variety of factors. The present review further suggests that the effect of phonemic awareness interventions can also vary based on the outcomes being reported, ranging from as small as g = .25 for sound deletion to g = .57 for sound segmentation. A summary of the effects observed in the present meta-analysis is included in Table 3.11 for ease of review. Measures of deletion were commonly seen to be those with the lowest observed gain relative to other phonemic awareness outcomes, which is somewhat expected given the more complex nature of the task (Yopp, 1988). Lower outcomes on this measure might be related to other factors such as working memory since children must segment the given word, identify the requested sound in the sequence, remove it from the word, and then blend the remainder of the word back together (or retain the rest of the sounds as a whole unit) in order to successfully complete the task.

Table 3.11

Summary of Intervention Effects by Phonemic Awareness Outcome

Outcome	Study *N*	Participant *N*	*g* [95% CI]
Composite	43	3,832	.51 [.38; .64]
Segmentation	41	2,909	.57 [.48; .67]
Blending	35	2,405	.34 [.22; .47]
Deletion	31	2,213	.25 [.12; .37]
FSID	21	1,306	.43 [.31; .54]

Subgroup analyses of the studies included in the present review should be interpreted with caution due to small sample sizes within each group but do provide some helpful information in interpreting the effects of phonemic awareness instruction with at-risk children. Due to the wide variety of sample sizes included in this analysis, sample size was binned to represent whether a study included less than 30 participants or more based on a simplification of the central limit theorem (Field, 2018). Because sample size was not significantly correlated with effect sizes for any of the measured outcomes, there can be greater confidence in the effects obtained from even some of the smaller studies included in this review (e.g., Kerins et al., 2010). In fact, nearly identical effects of phonemic awareness instruction were observed for both small and large samples, especially for segmentation ($Q = .06$, $p = .805$) and blending ($Q = < .05$, $p = .963$) outcomes. This information combined with the results of tests for small study effects suggests that educators can be more confident in the outcomes of smaller studies addressing phonemic awareness instruction, which are common in educational research (e.g., Frizelle et al., 2021).

Children from lower socioeconomic backgrounds tend to benefit from phonemic awareness instruction, with effect sizes ranging from small ($g = .289$) on measures of deletion to medium ($g = .529$) on those of segmentation. Regarding instructional timing, medium effects of phonemic awareness instruction were consistently observed across outcome measures for kindergarten children with the exception of the more developmentally complex deletion measure. This is consistent with the existing evidence that suggests many of the risk factors for reading failure are associated with entering school with depressed levels of informal literacy exposure and instruction (Kamhi & Catts, 2012). Although effects across outcomes varied by grade, interventions provided in kindergarten and first grade most often resulted in positive outcomes that straddled Cohen's (1988) relatively arbitrary distinction between what constitutes a small or medium effect. Because children entering school with poor phonological awareness skills tend to struggle throughout their formal schooling (Torgesen et al., 1994), intervention in these early grades is crucial to preventing reading failure.

The present meta-analysis also indicates that phonemic awareness instruction is not only helpful during the early elementary years but can also have significant effects on children's acquisition of basic reading skills well into the middle of the developmental period. Although the overall effect of phonemic awareness interventions does seem to decline as children age, large effects are still possible for children as old as 13 years as demonstrated by the work of Mirbazel et al. (2016), whose phonemic awareness instruction resulted in a 1.42 standard deviation gain over peers receiving business-as-usual instruction that did not include phonemic awareness training. Understanding the effects of delayed phonemic awareness instruction is critical, especially for students who have had minimal exposure to code-based instruction in the early grades for any number of reasons.

Interestingly, there was not generally a statistically significant difference between phonological and phonemic awareness instruction on phonemic awareness outcomes. Although data are presented above regarding both individual phonological and phonemic awareness instruction as well as either intervention with additional literacy instruction, statistically nonsignificant results remained across outcomes when these data were collapsed into only two groups having either a phonological or phonemic focus. The lack of statistical significance in the present data is likely an artifact of the scope of the present analysis: some amount of instruction at the phonemic level was present in every study in order it to be included in the review. There was, however, a significant difference between the two levels of intervention complexity with regard to composite outcomes. Phonemic awareness instruction had a large effect ($g = .814$) on composite outcomes compared to phonological awareness instruction ($Q = 6.01$, $p = .014$), perhaps suggesting that phonemic awareness instruction affects more widespread or generalized change. This finding is consistent with the previous literature suggesting that instruction should focus on the phonemic level of complexity as early as kindergarten (Schuele & Boudreau, 2008).

Of particularly good news for schools already burdened with staffing and resource shortages is the finding that phonemic awareness instruction can be effective when delivered by a wide variety of providers. Speech-language pathologists, teachers, and trained volunteers all demonstrated the ability to have at least a medium effect on phonemic awareness outcomes with at-risk children, though speech-language pathologists did demonstrate the largest overall effect on composite outcomes (g = .914). Furthermore, computer-based instruction was also found to be effective in training phonemic awareness across outcomes with the exception of those based on sound deletion (g = .035). For schools with the resources to do so, phonemic awareness instruction for many students might be able to be provided during supervised, computer-based instruction such as that used by Kartal et al. (2016) or Carson (2020) while reserving more intensive, staff-based instruction for children with the most severe needs. Such a system could be consistent with the existing RTI or MTSS model of tiered instruction being used in US schools to address student learning needs, where tier one instruction is provided to all students by classroom teachers (e.g., Carson et al., 2019), tier two instruction is provided through computer-based instruction (e.g., Chera & Wood, 2003), and tier three instruction is provided by a special education teacher or speech-language pathologist (e.g., Ukrainetz et al., 2009).

The effects associated with small and large group instruction are also continued good news for educators, with the present results confirming again that both small and large groups can be very effective instructional arrangements. Although a statistically significant difference between arrangements was not detected in the present review, this may also be the result of small samples not being large enough to detect different effects. All three distinct instructional arrangements coded in the present review (individual, small group, large group) were capable of delivering at least medium effects on phonemic awareness outcomes. Although Ehri et al. (2001) found a large effect associated with small group instruction (d = 1.38) compared to medium effects for both individual (d = .60) and classroom-based instruction (d = .67), the present meta-analysis found comparable effects between the three instructional arrangements across outcomes. It should be noted, however, that small and large group arrangements combined (n = 75) were reported more frequently in the accumulated literature on work with at-risk children than individual instruction (n = 63).

Consistent with the findings of previous reviews, more intervention does not always result in better outcomes (Ehri et al., 2001; Suggate, 2016). These data suggest that it is the components of the intervention being provided rather than the length of the intervention itself that are likely more important in making inferences about the effects of phonemic awareness interventions with specific children. Future research should account for the productivity of individual sessions rather than just the length, frequency, and duration of intervention sessions in order to better understand the specific factors responsible for affecting change in children's phonemic awareness abilities (Frizelle et al., 2021). Because intervention research to date has largely reported how long interventions have been provided to children rather than the productivity of each session, it is possible that these findings might be used to suggest that children require a specific prescribed amount of instruction in phonemic awareness before other skills can be addressed, which would be an inappropriate but understandable conclusion based on the data presently available to consumers. Although Ehri et al. (2001) found significant differences favoring phonemic awareness instruction provided for five to 10 hours over other amounts, Suggate (2016) and the present review agree in finding that the amount of intervention provided is not significantly related to intervention outcomes.

Regarding the components of interventions, the data continue to suggest that graphemes should be incorporated to the extent appropriate for children's knowledge of sound-symbol correspondence (Spector, 1995). Specifically, the present meta-analysis continues to support the meaningful inclusion of graphemes into phonemic awareness instruction based on statistically significant differences in gains associated with their inclusion observed for composite outcomes ($Q = 4.84$, $p = .028$). Larger gains related to the inclusion of graphemes were also demonstrated on outcomes of segmentation ($g = .594$), blending ($g = .337$), and deletion ($g = .253$) as well, though not to a statistically significant level. Educators should incorporate graphemes into phonemic instruction when working on segmentation and blending skills especially, as this facilitates orthographic mapping and direct opportunities to practice decoding in a meaningful context (Schuele & Boudreau, 2008).

Limitations

Limitations of the present meta-analysis include the decision to utilize multiple outcomes per included study rather than aggregating within-study effects, which is likely to have biased the estimated effects (Field, 2018). Because multiple outcomes were investigated that measure different theoretically-related skills (Schuele & Boudreau, 2008; Yopp, 1988), this method was preferred to avoid introducing additional heterogeneity into the data by clustering outcomes together. Despite the likelihood of biased estimates due to the methods used, the results of the present meta-analysis are consistent with previous meta-analyses of phonemic awareness instruction (Ehri et al., 2001; Suggate, 2016). The use of a random-effects model also allows for generalization of the present analyses to other work in the area, especially when combined with the results of tests for small-study effects (Schwarzer et al., 2015). Finally, although data were extracted for the comparison of phonemic awareness interventions to modified or enhanced phonemic awareness interventions, relevant variables exploring the differences between standard phonemic awareness interventions and those that were augmented or modified in some way (e.g., Anderson, 2010; Hatcher et al., 2006) were not collected. Future research should specifically investigate the components incorporated into phonemic awareness interventions to help educators identify the most efficacious methods of providing such instruction.

Implications

The present work continues to support the use of phonemic awareness instruction with children at risk of reading failure, building upon previous analyses conducted by Ehri et al. (2001) and Suggate (2016). Phonemic awareness instruction is consistently associated with positive effects for children at risk of reading failure when provided by a wide variety of individuals, including paraprofessionals (e.g., Lane et al., 2007). This instruction can also be effective across a variety of instructional arrangements, giving educators great leeway in meeting individual students' needs while being mindful of often limited resources. Educators should also be mindful that substantial gains can be made in a short amount of time with children at-risk of reading failure (e.g., Chera & Wood, 2003). Although interventions provided in kindergarten and first grade remain associated with the most consistent gains, the development of adequate phonemic awareness can be achieved with older students as well—especially regarding phonemic segmentation and blending.

CHAPTER FOUR

Contextualized Phonemic Awareness Instruction with At-Risk Readers

Because phonological awareness is a pre-reading skill, its utility in contributing to models of reading achievement is strongest *before* children are reading; after children have begun reading, screenings and assessments should focus on the skills utilized in reading itself rather than pre-reading skills such as phonemic awareness. As children become fluent word decoders, phonemic awareness instruction has less of an impact on reading performance and contributes less to models of reading achievement (Kamhi & Catts, 2012). While many studies have been conducted on the outcomes of phonemic awareness training during the emergent literacy period and early school years, phonemic awareness instruction is largely ignored past the first grade (Ehri, et al., 2001; Suggate, 2016). In applied practice, however, it is likely that some children may enter second grade with insufficient phonemic awareness for a variety of reasons, including poor intervention referral practices or even inadequately designed instruction (Chaney, 1990). Even in the most recent meta-analysis including outcomes for phonemic awareness instruction, only four of 17 studies investigated outcomes in the second grade or higher and three of those four focused on their administration via computer programs rather than highly qualified instructional staff (Suggate, 2016). Especially given the emerging impact of covid-19 on the U.S. education system, it is very likely in the coming years that students will enter the second grade with incomplete phonemic awareness training that translates to inadequate word decoding and inhibits their reading comprehension development. More information is needed on the effects of phonemic awareness instruction when provided to older children.

For children receiving response to intervention (RTI) services to address deficient phonemic awareness, target outcomes should include the segmentation of words into their constituent sounds, as this skill is most directly related to reading success (Schuele & Boudreau, 2008). When children are able to segment words into their individual phonemes, they should then be capable of combining this skill with their print knowledge to decode unknown words. Because of this combined interest in phonemic segmentation and print knowledge, it is generally recommended that phonemic awareness interventions include a component designed to develop children's understanding of sound-symbol correspondence or, rather, the relationship between phonemes and their corresponding graphemes (Kamhi & Catts, 2012; Schuele & Boudreau, 2008; Ukrainetz, 2015). Until children are able to accurately and fluently segment words into their individual phonemes, intervention should continue on a data-driven basis to support the development of these skills. While the development of phonological awareness is "quasi-parallel" (McNeill et al., 2017, p. 304), meaning that there is a rough developmental sequence but mastery of one skill is not necessarily a pre-requisite for mastery of skills later in the sequence, there is a general order of recommended instruction for the development of phonemic awareness (Schuele & Boudreau, 2008).

Broadly speaking, phonemic awareness instruction begins with the segmentation of words into their onset and rime components followed by the segmentation of only the initial or final sounds. After that, instruction tends to focus on blending sounds to form words and then segmenting spoken words into their individual sounds. The last skill often targeted in phonemic awareness instruction is the deletion of individual sounds within a word, however, segmentation of words into phonemes is the skill most directly linked to improved reading outcomes (Schuele & Boudreau, 2008). For students enrolled in preventative instructional services (i.e., RTI), the culmination of phonemic awareness training should be accurate and fluent word decoding if those interventions focus on developing their phonemic awareness while also ensuring accurate, reliable, and fluent phoneme-grapheme correspondence (Suggate, 2016). This can readily be accomplished through print-referencing and joint engagement in storybook reading activities, which maximizes the amount of time spent in phonemic awareness instruction while also ensuring intentional exposure to the printed word (Justice & Ezell, 2004; Piasta et al., 2012).

Contextualized instruction, or the embedding of instruction into the context(s) to which it is most relevant, can be an effective method of programming for generalization (Gillam et al., 2012; Peterson, 2011; Ukrainetz, 2015). Decontextualized instruction, or that occurring in isolation without a direct link to the target activity, is efficient in providing high numbers of trials to perform a given skill but can suffer from poor generalization to meaningful contexts (e.g., Fuchs & Fuchs, 2002; Gillam et al., 2012). While a review of the literature suggests that decontextualized instruction is the most common modality of instruction for speech-language pathologists, the embedding of instruction within reading activities is becoming more popular in recent years (Ukrainetz, 2015). By embedding interventions designed to address spoken and/or written language into the context of narrative storybooks, the effects of instruction are more naturally generalized to the same activities where they occur outside of the therapeutic context

(Ukrainetz, 2015), which is especially important when working with at-risk populations. Because the whole point of RTI and related interventions is to affect change outside of the instructional or therapeutic environment, speech-language pathologists and other professionals involved in the education of children should consider the use of contextualized instruction to promote skill transfer across contexts within similar activities. To date, however, there is conflicting research on the relative effectiveness of contextualized and decontextualized instruction for children with language disorders, with some skills showing greater gains when taught contextually and others demonstrating greater improvement without being embedded within natural contexts (e.g., Gillam et al., 1995; Gillam et al., 2012; Swanson et al., 2005).

Although it is clear that phonemic awareness instruction is effective with most children (Ehri, et al., 2001; Suggate, 2016), the present study seeks to demonstrate the effectiveness of these interventions when they are specifically embedded within a storybook reading activity (i.e., provided within a contextualized intervention framework). Furthermore, despite knowing that phonemic awareness interventions are efficacious, there continues to be a lack of data on their effectiveness when implemented in applied settings, especially with children older than first grade. The present study seeks to fill that gap by implementing contextualized phonemic awareness instruction in a public school with regard for the naturally existing constraints of that setting.

Method

This research was approved by the Baylor University Institutional Review Board (IRB) and collaborating school district's administration prior to the occurrence of any study activities.

Inclusion and Exclusion Criteria

The principal investigator collaborated with the administration of a rural, Title 1 elementary school in Central Texas to complete this research and contribute to a review of the school's existing RTI procedures for children at risk of reading failure. In order to be considered for inclusion in this study, children had to be (a) enrolled in either virtual or in-person instruction with the collaborating elementary school at the beginning of the 2020-21 school year, (b) not currently receiving special education or section 504 supports and services, and (c) previously classified as "at risk" of developing a reading disorder as discussed below.

Children enrolled in the second grade at the beginning of the 2020-21 school year were previously administered the STAR Early Literacy Assessment (Renaissance Learning, 2017) as a screening measure of reading achievement and the school uses student performance on this instrument as an indicator of risk for the presence of a reading disorder. Based on a student's overall performance on the STAR Early Literacy Assessment, they are classified as either meeting benchmark standards or in need of monitoring, intervention, or urgent intervention. Student performance on the STAR Early Literacy Assessment is also reported in a variety of domains, including their understanding of the alphabetic principle, concept of words, visual discrimination, phonemic awareness, phonics, structural analysis, vocabulary, sentence-level comprehension, paragraph-level comprehension, and early numeracy. Students receive scores in each of these domains based on their mastery of related skills (listed in Table 4.1) as well as an overall score, which is then used to identify them as an early emergent, late emergent, transitional, or probable reader.

Table 4.1

STAR Early Literacy Domains and Associated Skills

STAR Domain	Skill Set
Alphabetic Principle	Alphabetic Knowledge
	Alphabetic Sequence
	Letter Sounds
Concept of Word	Letters and Words
	Word Borders
	Word Length
Paragraph-level Comprehension	
Phonemic Awareness	Blending Phonemes
	Blending Word Parts
	Consonant Blends
	Initial and Final Phonemes
	Medial Phoneme Discrimination
	Phoneme Isolation and Manipulation
	Phoneme Segmentation
	Rhyming and Word Families
Phonics	Consonant Blends
	Consonant Digraphs
	Final Consonant Sounds
	Initial Consonant Sounds
	Long and Variant Vowel Sounds
	Other Vowel Sounds
	Short Vowel Sounds
	Sound-symbol Correspondence (Consonants)
	Sound-symbol Correspondence (Vowels)
	Word Building
	Word Families and Rhyming
Sentence-level Comprehension	
Structural Analysis	Compound Words
	Syllabification
	Words with Affixes
Visual Discrimination	Identification and Word Matching
	Letters
Vocabulary	Antonyms
	Synonyms
	Word Facility

When collaborating with school administrators to identify students who would benefit the most from the intervention being investigated in the present study, the principal investigator encouraged them to prioritize those students with the lowest phonemic awareness scores, especially those whose performance in other domains was markedly higher. Because of school closures associated with the COVID-19 pandemic, school administrators also indicated that they would like to prioritize referrals of students who had not actively engaged in distance learning during the Spring 2020 term and were already struggling to acquire basic reading skills. Based on the eligibility criteria discussed above, school administrators identified 50 children as being at-risk of reading failure at the beginning of their second-grade year and were thus eligible for consideration to participate in this study.

Data Collection

Several measures were used to evaluate participant eligibility for participation and monitor their progress throughout the study. The *Kaufman Brief Intelligence Test-Second Edition* (KBIT-2; Kaufman & Kaufman, 2004) is an individually administered, norm-referenced assessment of both verbal and nonverbal intelligence appropriate for individuals between the ages of four and 90 years old. Participants were administered the full KBIT-2, which includes the Verbal Knowledge, Riddles, and Matrices subtests. Raw scores on the Verbal Knowledge and Riddles subtests are combined to form the KBIT-2's Verbal composite score while the Matrices subtest is used as the sole measure comprising the Nonverbal composite. An individual's standard scores on the Verbal and Nonverbal composites are then used to calculate an overall estimate of their intellectual ability represented by the IQ composite score. For all three composites, scores between 85 and

115 are considered to be within the average range for an individual's age. The authors report internal consistency estimates of .90, .86, and .92 for children aged four to 18 years on the Verbal, Nonverbal, and IQ composites, respectively.

The *Comprehensive Test of Phonological Processing-Second Edition* (CTOPP-2; (Wagner et al., 2013) is an individually administered, norm-referenced assessment of phonological processing abilities appropriate for individuals between the ages of four and 24 years old. For the ages covered by the current study, the CTOPP-2 consists of seven core and two supplemental subtests. For individual subtests, scores of eight to twelve are considered to represent the average range of ability while scores of 90 to 110 are considered average for each composite. An overview of each subtest is provided in Table 4.2 along with how those subtests are combined to form the CTOPP-2's four composites. The authors report internal consistency estimates of .77 to 92 for the full age range of the normative sample; all composite internal consistency estimates were greater than .85.

Table 4.2

CTOPP-2 Composite and Subtest Descriptions

Measure	Description
Phonological Awareness	Represents the awareness of and access to language's phonological structure
Elision	Measures the ability to delete phonemes in words
Blending Words	Measures the ability to blend phonemes into words
Phoneme Isolation	Measures the ability to identify phonemes in words
Phonological Memory	Represents the ability to code phonological information in working memory
Memory for Digits	Measures the ability to repeat sequences of numbers
Nonword Repetition	Measures the ability to repeat nonwords
Rapid Symbolic Naming	Represents the ability to retrieve phonological information from long-term memory
Rapid Digit Naming	Measures the speed with which one can name single numbers
Rapid Letter Naming	Measures the speed with which one can name single letters
Alt. Phonological Awareness	Represents the awareness of and access to language's phonological structure without the influence of previous word knowledge
Blending Nonwords	Measures the ability to blend phonemes into nonwords
Segmenting Nonwords	Measures the ability to segment nonwords into phonemes

The Dynamic Indicators of Basic Early Literacy Skills-Eighth Edition (DIBELS-8; University of Oregon, 2020) is a standardized assessment of early literacy skills that can be used for progress monitoring children's progress toward acquiring appropriate literacy skills through eighth grade. The DIBELS-8 includes both benchmarks designed to be administered at the beginning, middle, and end of an academic year as well as progress monitoring forms that may be administered throughout the course of an intervention. Four DIBELS-8 measures are recommended for use with second grade students, including Nonsense Word Fluency (NWF), Word Reading Fluency (WRF), Oral Reading Fluency (ORF), and Maze. Because the focus of this project is on the acquisition of phonemic awareness, the Phoneme Segmentation Fluency (PSF) measure was also included as both a benchmark and progress monitoring tool despite grade-level comparison data being unavailable. In addition to 20 progress monitoring forms for each measure, the DIBELS-8 also has three benchmark assessments for each measure, one for the beginning, middle, and end of the year.

PSF is a standardized, timed assessment of a child's ability to fluently segment words into individual phonemes. On this measure, examiners verbally present a series of words that the child is asked to verbally segment over the course of sixty seconds. Students earn points based on the number of individual sounds correctly produced within the time allotted. NWF is a standardized, timed assessment of a child's ability to decode written text without relying on known vocabulary and sight words. On this measure, children are presented with a list of written nonsense words and are asked to read as many of them as they can either in part (e.g. r-a-l) or in whole (e.g., "ral") in 60 seconds. WRF is a standardized, timed assessment of a child's ability to read real words in isolation. On this measure, children are presented with a list of written real words and asked to read as many of them as they can in 60 seconds. Unlike on the NWF measure, children must read the whole word correctly on WRF to receive credit for the word. While WRF is a measure of a child's ability to read words in isolation, ORF is a measure of their ability to read accurately and fluently in connected text. On this measure, the child is presented with a passage and given one minute to read out loud as much of the passage as they can. Finally, Maze is a measure of a child's ability to understand what they are reading. On this measure, every seventh word is removed and replaced with three options as to what the word should be based on the content of the passage. As they read, children are asked to circle which of the three options is correct for each missing word they encounter within three minutes. A summary of scoring procedures for the

DIBELS-8 is included in Table 4.3, where an adjusted Maze scoring procedure is described to account for guessing.

Table 4.3

DIBELS-8 Scoring Procedures

Measure	Scoring Procedure
PSF	The number of correct phonemes produced correctly in 60 seconds.
NWF-CLS	The number of correct letter sounds produced in 60 seconds.
NWF-WRC	The number of nonsense words read correctly in 60 seconds.
ORF-ACC	The number of words read correctly divided by the total number of words read in 60 seconds multiplied by 100.
ORF-WRC	The number of words read correctly in 60 seconds.
WRF	The number of words read correctly in 60 seconds.
Maze Adjusted	The number of total words identified correctly minus one-half of the number of words attempted in three minutes.

The principal investigator received a list of students who met the study's broad inclusion criteria and collaborated with school personnel to call the caregivers of these children in an order recommended by school staff. Three attempts at contacting a student's caregiver were made before removing them from consideration and attempting to contact the next potential participant on the list. In total, the caregivers for eight children were contacted and informed consent was obtained for four to participate. The caregivers of three children did not respond to attempts to contact them and the caregiver of the remaining potential participant expressed interest in the study but did not return the consent form after three attempts to obtain it.

After informed consent was obtained, the principal investigator administered the KBIT-2 to screen potential participants' verbal and nonverbal intelligence. Participants whose scores on the KBIT-2 fell below 70 were to be excluded from further participation because tiered interventions such as the one being investigated presently are unlikely to be sufficient in meeting their learning needs (Al Otaiba & Fuchs, 2002). None of the children screened were excluded on the basis of their KBIT-2 scores, which are provided in Table 4.4.

Table 4.4

KBIT-2 Scores by Participant

Measure	Participant 1	Participant 2	Participant 3	Participant 4
Nonverbal	108	88	92	88
Verbal	119	89	97	97
IQ Composite	116	88	94	91

Participants were then administered the CTOPP-2 to establish a baseline of their phonological awareness abilities in comparison to a nationally representative sample of their peers. Participants whose scores on the CTOPP-2 fell within or above the average range—especially on phonological or alternate phonological awareness measures—were to be excluded from further participation due to a lack of need for the intervention being used in the current study. None of the participants screened were excluded on the basis of their CTOPP-2 scores, which are reported in Table 4.5.

Table 4.5

CTOPP-2 Pretest Scores by Participant

Measure	Participant 1	Participant 2	Participant 3	Participant 4
Subtests				
Elision	7	4	6	7
Blending Words	9	7	8	9
Phoneme Isolation	5	4	5	7
Memory for Digits	10	4	5	5
Nonword Repetition	10	5	8	5
Rapid Digit Naming	9	9	9	10
Rapid Letter Naming	8	9	7	10
Blending Nonwords	7	6	7	8
Segmenting Nonwords	4	7	9	7
Composites				
Phonological Awareness	82	69	77	86
Phonological Memory	101	67	79	70
Rapid Symbolic Naming	92	95	88	101
Alt. Phonological Awareness	73	79	88	85

Once participants' eligibility for participation was established on the basis of their KBIT-2 and CTOPP-2 scores, pre-intervention baseline testing was completed with the DIBELS-8 beginning of year benchmark form. Pre-intervention baseline performance on the DIBELS-8 is reported for each participant in Table 4.6 with risk classifications reported in Table 4.7. All four participants enrolled in the study were aged seven or eight years at the beginning of the intervention (*M* = 7.88 years, *SD* = .14 years), were male, and were enrolled in a Title 1 school but specific information regarding their individual socioeconomic status (SES) was not requested. None of the four students had previously repeated a grade nor were they enrolled previously in tier three interventions. Caregivers reported no sensory impairments or other medical conditions that would likely have affected their participation in this study. Participants one and two were enrolled in virtual learning at the beginning of the study while participants three and four were enrolled in in-person learning during their participation. Due to delays in obtaining informed consent from interested caregivers, participants three and four were unable to begin the intervention until approximately one month after participant one and two.

Table 4.6

DIBELS-8 Pretest Scores by Participant

Measure	Participant 1	Participant 2	Participant 3	Participant 4
PSF	16	16	46	55
NWF-CLS	71	36	59	40
NWF-WRC	21	7	12	8
ORF-ACC	94.34	91.35	88.88	78.57
ORF-WRC	100	74	48	22
WRF	47	15	28.2	12
Maze Adjusted	4	0	2	2
Composite[a]	360.38	329.64	331.63	311.36

[a] Composite scores were calculated using a weighted formula provided by the University of Oregon (2020) and do not include these children's performance on the PSF measure.

Table 4.7

DIBELS-8 Pretest Risk Classifications by Participant

Measure	Participant 1	Participant 2	Participant 3	Participant 4
PSF[a]	At Risk	At Risk	Min. Risk	Min. Risk
NWF-CLS	Min. Risk	At Risk	Min. Risk	Some Risk
NWF-WRC	Min. Risk	At Risk	Some Risk	At Risk
ORF-ACC	Min. Risk	Some Risk	Some Risk	At Risk
ORF-WRC	Min. Risk	Min. Risk	Some Risk	At Risk
WRF	Min. Risk	At Risk	Min. Risk	At Risk
Maze Adjusted	Some Risk	At Risk	At Risk	At Risk
Composite	Min. Risk	Min. Risk	Min. Risk	At Risk

[a] PSF risk classifications are reported based on the recommendations for children at the end of their first grade year.

Experimental Design

Because the intervention procedures used during RTI are focused on affecting change within individuals, a single case research design was preferred over a group design. Specifically, a multiple baseline design across behaviors was used to evaluate the effectiveness of the intervention on different phonemic awareness skills within individual participants. This design was then replicated across multiple participants. Multiple baseline research designs evaluate the effectiveness of an intervention by introducing it at different points in time (Kazdin, 2011). Specifically, multiple baseline designs that operate across behaviors introduce the intervention to different behaviors (i.e., skills) within the same individual at different points in time. By staggering the introduction of the intervention, the aim is to demonstrate that the behaviors change when and, ideally, only when the intervention is introduced, thus suggesting that the intervention drove the change in performance. In the present study, four skills were initially selected to be targeted: counting the number of phonemes in words, identifying the first phoneme in words (FSID), segmenting words into phonemes, and blending phonemes into words. When working with participants one and two, absences obfuscated visual analyses of the data and there were concerns that four skills may not be sufficient for clear replication of the intervention's effects. For this reason, a fifth skill, deleting phonemes from words, was added for participants three and four.

Because baseline (i.e., pre-intervention) data are used to establish current performance and predict future performance (Kazdin, 2011), relatively stable baseline performance is needed prior to the introduction of the intervention. For all participants, a minimum of five sessions (including testing sessions) were planned to obtain stable baseline performance before beginning the intervention on the first behavior (phoneme counting). If a participant demonstrated high baseline levels of performance on a skill in the sequence, it was omitted from the intervention sequence. For each participant, once performance on all skills was judged to be stable in that a clear level and trend (or lack thereof) were observed, the principal investigator introduced the intervention. After that, intervention continued on that same skill until the participant demonstrated a performance of 90% accuracy or greater on probe data collected at the beginning of each session over three consecutive sessions. At that time, the intervention was introduced to the next skill in the following sequence: first phoneme identification, phoneme segmentation, phoneme blending, or phoneme deletion. Because phonemic awareness is a pre-literacy skill, high levels of accuracy and fluency are needed so that students can transition seamlessly from phonologically decoding written words sound-by-sound towards decoding spelling patterns (Kamhi & Catts, 2012). The mastery criterion of correctly performing a target skill with 90% accuracy over three consecutive sessions before moving onto the next targeted skill was selected because of the need for independent and fluent performance of each skill by the end of the intervention.

Because each participant served as their own control in this multiple baseline study across behaviors study, visual analyses were conducted as the primary method of analysis. To support and quantify these analyses, statistical analyses were conducted to investigate the stability of performance at baseline and calculate appropriate quantitative representations of the intervention's effect based on the work of Tarlow (2017). These analyses and calculations involved the calculation of baseline trends and, if necessary, correction prior to the calculation of an effect size for each skill. Visual inspections were conducted in accordance with Kazdin (2011) regarding documentation of magnitude, rate, and overall patterns of change.

Procedures

The principal investigator scheduled a single standing appointment with the guardian of participants one and two, who were related, four days a week at a time convenient for them. Sessions for participants one and two were conducted via teleconferencing with the principal investigator using a document camera to share the storybook and manipulatives with the participants. For participants three and four, sessions were scheduled at a mutually agreeable time with their classroom teachers four days weekly and no technology was incorporated into their sessions with the exception of video recording equipment. All sessions were recorded via the same teleconferencing software for the purpose of collecting interobserver agreement and treatment fidelity data, which are discussed below. Each individual participant's sessions were scheduled for thirty minutes of individual instruction regardless of the delivery modality.

Each session during both the baseline and intervention phases of the study began with the principal investigator collecting probe data on all skills addressed during the course of the intervention. For each skill (counting the number of phonemes in words, segmenting words into phonemes, blending phonemes into words, identifying the first phoneme in words [FSID], and deleting the first phoneme in words), the principal investigator presented ten trials to the participant to elicit a response without feedback. Stimuli for these probes were selected from randomly selected DIBELS-8 PSF progress monitoring forms using a random number generator. That is, a DIBELS-8 PSF progress monitoring form was randomly selected and then the stimuli from that form were also randomly selected. Stimuli themselves were randomly selected because the standardized administration procedures of the DIBELS-8 require individuals to administer the target words in order, where the stimuli themselves are ordered to become more complex as the child progresses through the form. Randomizing the difficulty of the stimuli presented to each participant was intended to serve as a less biased measure of their performance than only presenting them with the monosyllabic, less complex items at the beginning of each PSF form.

After probe data were collected, the principal investigator informed or reminded the participant of the skill being targeted during each day's session. Often, this was accomplished by asking the participant "do you remember what we're working on today?". Immediately following that, the participant was provided with eight colored wooden blocks for use as manipulatives during instruction. The principal investigator then read one-to-two pages of the week's platform storybook before engaging the participant in discrete trial training of the day's targeted phonemic awareness skill. This process of reading one or two pages of text followed by discrete trial training was repeated for the remainder of each day's session. Stimuli for use in discrete trial training were pulled from the pages read with the participant or selected by Goldsworthy (2012), which was also the source used to identify books for use in the study.

Texts were selected for use in the present study based on those included in Goldsworthy (2012), who provided selected stimuli for eight classic children's stories, including *Goldilocks and the Three Bears, Jack and the Beanstalk, Little Red Riding Hood, Rumpelstiltskin, Sleeping Beauty, Snow White and the Seven Dwarfs, The Gingerbread Boy,* and *Three Billy Goats Gruff*. Of these texts, the principal investigator utilized several adaptations that were told from a different perspective than the classic works so as to maintain engagement while providing a familiar overall narrative structure for the children to use for background knowledge. A list of specific books used during the course of the present study is provided in Table 4.8 and all but one of the books were from the same series of works to promote cohesion from week to week. Generally, each book was read in its entirety over the course of three days with the fourth day's instructional session being used as a quick review of the entire work. Constraints in the school's calendar resulting from the time intervention was able to be started with participants three and four prevented the use of all nine planned books with each participant. Additionally, both participants three and four requested the use of the Braun (2012) text due to its availability in the office space where instruction was provided during the last week of the intervention before the school's holiday break. Data were not specifically collected on which books were used with each participant but each of the texts listed in Table 4.8 was used with at least two participants during the course of the intervention. Participants were exposed to different platform books based on when they were enrolled in the study, all of which are included in Table 4.8. The same books were used for participants one and two while a different combination of books was used for participants three and four.

Table 4.8

Platform Books Utilized in Intervention

Title	Narrator
Believe me, Goldilocks rocks! (Loewen, 2012)	Baby Bear
Frankly, I'd rather spin myself a new name! (Gunderson, 2016)	Rumpelstiltskin
Honestly, Red Riding Hood was rotten! (Shaskan, 2012)	The Wolf
Listen, my bridge is so cool! (Loewen, 2018)	The Troll
No lie, pigs (and their houses) can fly! (Gunderson, 2016)	The Wolf
Seriously, Snow White was so forgetful! (Loewen, 2013)	The Dwarves
Snowmen at Night (Buehner, 2002)	N/A
Trust me, Hansel and Gretel are sweet! (Loewen, 2016)	The Witch
Trust me, Jack's beanstalk stinks! (Braun, 2012)	The Giant

The bulk of each instructional session was focused on discrete trial training via joint reading of the week's platform text. Although manipulatives were planned to keep participants occupied and maintain attention during sessions, this structure was not available to participants one and two via distance learning and was unnecessary for participants three and four during in-person instruction. However, manipulatives in the form of colored blocks were provided as an instructional material for all participants. For all four participants, the principal investigator read one to two pages of the target book and then engaged them in discrete trial training using a combination of stimuli pulled from Goldsworthy (2012) and stimuli selected from the pages that were just read. Stimuli pulled from Goldsworthy (2012) were almost exclusively words relevant to the content of each story (e.g., Goldilocks, granny, beanstalk) while words pulled from the texts themselves varied between both function and content words; stimulus difficulty (e.g., word length, syllable structure, presence of clusters) was adjusted continuously both within and across sessions to meet participant learning needs. The pace of discrete trial training was also adjusted to meet the needs of each participant in order to maintain high levels of attention and responding. Productivity data are reported later and it is evident that sessions were not equally productive across participants.

Data Collection

At the beginning of each session, the principal investigator collected these probe data by administering ten trials per skill without providing feedback to the participant.

Data were collected on each trial administered during sessions and participant responses were recorded with respect to the level of independence exhibited; a legend is provided in Table 4.9. A most-to-least prompting hierarchy was used to facilitate early and consistent access to correct performance and reinforcement.

Table 4.9

Prompting Hierarchy

Prompt	Label	Description
+	Correct	The participant independently performed the task correctly.
IV	Indirect Verbal or Visual	The participant performed the task correctly in response to repeated presentation of the question or the PI using a verbal or visual cue to indicate "are you sure?".
FC	Forced Choice	The participant correctly performed the task when presented with two response options.
IM+	Imitation with Time Delay	The participant correctly performed the task when the PI modeled the correct response, waited five seconds, and repeated the trial.
IM	Imitation	The participant correctly performed the task in direct imitation of the PI.
-	Incorrect	The participant was unable to correctly perform the task even with direct imitation of the PI

Progress monitoring using the complete DIBELS-8 progress monitoring system was conducted every other week in accordance with recommendations that progress monitoring occur no more frequently than every two weeks for children in kindergarten through third grade (University of Oregon, 2020). For each of the DIBELS-8 progress monitoring measures, individual forms were randomly selected for administration every other week over the course of the intervention. During these sessions, daily probe data were collected first as described above and then each of the DIBELS-8 measures used in this study were administered before engaging in regularly scheduled intervention procedures with the remaining time.

Interobserver Agreement and Treatment Fidelity

Two advanced graduate students contributed to the collection of interobserver agreement (IOA) and treatment fidelity data. For IOA data collection, the principal investigator randomly selected approximately 30% of the highest session count obtained (28) for review. That is, nine sessions for each participant were randomly selected for review. One advanced graduate student with previous experience teaching phonological awareness and familiarity with the skills being measured viewed video recordings of daily probe data collection and independently counted the number of correct performances. IOA was measured by percentage of agreement between raters per session. Probe data subject to IOA rating were collected on four skills for participants one and two while a fifth skill was added for participants three and four. This means that sessions for participants one and two contained four opportunities for agreement while sessions for participants three and four contained five opportunities. IOA data for all four participants are reported in Table 4.10. Unfortunately, technical difficulties in recording sessions resulted in audio issues that primarily affected participant three. The two sessions in which complete audio was unavailable for review were counted as missing data and partial data are not included in Table 4.10. Despite these technical issues, a minimum of 20% of sessions for each participant were reviewed as suggested by WWC (2020). We collected IOA data for 47% of sessions for participants one and two, 25% of sessions for participant three, and 31% of sessions for participant four. For daily probe data, the mean percentage of agreement for each participant across the duration of the study met the 80% threshold recommended in the literature (Hartmann et al., 2004).

Table 4.10

Interobserver Agreement Data

	Sessions				Percentage Agreement				
Participant	Total	Reviewed	Complete	Percentage	Min.	*Med*	*Mn*	*SD*	Max.
1	19	9	9	47	75	100	97.22	8.33	100
2	19	9	9	47	75	100	91.67	12.50	100
3	28	9	7	25	100	100	100	0	100
4	28	9	9	31	80	100	95.56	8.82	100

The principal investigator selected approximately 30% of the highest session count obtained (28) for review by another advanced graduate student for the purpose of collecting treatment fidelity data. That is, nine sessions were randomly selected for each participant to be reviewed for treatment fidelity purposes and were selected separately from those selected for IOA data collection. Treatment data for each reviewed session were collected in the areas of session organization, probe data collection procedures, activity presentation, and discrete trial training procedures. A list of criteria used for collection of treatment fidelity data are included in Table 4.11, the number of observed sessions are reported in Table 4.12, and treatment fidelity data themselves are reported in Table 4.13. Treatment fidelity was measured based on the percentage of steps completed correctly within each session for each participant. These data were collected for 47% of sessions for participants one and two and 32% of sessions for participants three and four. The average treatment fidelity for participants one and three was 100% with no variability in procedures observed. For participants two and four, the average treatment fidelity was 99.5% with a range of 98 – 100%. For these two participants, the sole issue observed with deviations from treatment procedures was related to the principal investigator not telling the participant what they would be working on during the session. Because phoneme deletion was not addressed with participants one and two, criteria regarding the collection of their probe data were not included in treatment fidelity assessments for their sessions.

Table 4.11

Treatment Fidelity Checklist

Content Area	Criterion
Session Organization	The first activity (excluding DIBELS-8 testing) is probe data collection. The second activity (excluding DIBELS-8 testing) is reading the storybook. The third activity (embedded within activity two) is discrete trial training.
Probe Data Collection	Probe data are collected for phoneme counting. Probe data are collected for first sound identification. Probe data are collected for phoneme blending. Probe data are collected for phoneme segmenting. Probe data are collected for phoneme deletion. 10 probe trials are conducted for phoneme counting. 10 probe trials are conducted for first sound identification. 10 probe trials are conducted for phoneme blending. 10 probe trials are conducted for phoneme segmenting. 10 probe trials are conducted for phoneme deletion No corrective feedback is provided during probe data collection.
Activity Presentation	The PI tells the participant what they will be working on during the session. The PI reads the storybook with the participant. The PI intersperses DTT within the reading. The PI provides corrective feedback in response to participant errors. The PI maintains participant engagement during the session.
DTT Procedures	The PI provides a minimum of 20 learning trials during the session. The PI provides up to three seconds for participants to generate an independent response. The PI provides corrective feedback to all incorrect responses that are not self-corrected by the participant.

Table 4.12

Treatment Fidelity Session Counts

Participant	Total	Reviewed	Complete	Percentage
1	19	9	9	47
2	19	9	9	47
3	28	9	9	32
4	28	9	9	32

Table 4.13

Treatment Fidelity Data

Area by Participant	Percentage				
	Min.	*Med*	*Mn*	*SD*	Max.
1					
Session Organization	100	100	100	0	100
Probe Data Collection	100	100	100	0	100
Activity Presentation	100	100	100	0	100
DTT Procedures	100	100	100	0	100
2					
Session Organization	100	100	100	0	100
Probe Data Collection	100	100	100	0	100
Activity Presentation	80	100	98	7	100
DTT Procedures	100	100	100	0	100
3					
Session Organization	100	100	100	0	100
Probe Data Collection	100	100	100	0	100
Activity Presentation	100	100	100	0	100
DTT Procedures	100	100	100	0	100
4					
Session Organization	100	100	100	0	100
Probe Data Collection	100	100	100	0	100
Activity Presentation	80	100	98	7	100
DTT Procedures	100	100	100	0	100

Results

Recruitment and Participant Flow

Four participants were recruited to participate in this study. During the course of the study, participants one and two demonstrated considerable attendance and participation issues that impacted intervention administration and data collection. To address these issues, the principal investigator scheduled a daily standing appointment and sent a reminder text message to their guardian approximately thirty minutes before each scheduled session per their request. For each session in which the participants did not log on within 15 minutes of their scheduled start time, the principal investigator called the guardian to remind them again of the scheduled appointment and ask if another time would work better for them. On several occasions, sessions were rescheduled on the same day or made up within the same week. Despite these efforts, participants one and two each accumulated 14 missed sessions that were not made up over the course of 33 attempted sessions, missing approximately 42% of their scheduled intervention. Because these two participants missed over a third of their planned intervention sessions, they were administratively removed from the study. Participant three missed two of 30 planned sessions due to scheduling conflicts on the part of the principal investigator, one of which was able to be made up. Participant four missed three of 31 planned sessions, one of which was a scheduling conflict on the part of the principal investigator and two of which were related to the participant being sick. The two sessions he missed due to illness occurred on the last two days of planned intervention and were unable to be rescheduled due to a two week break in the school's instructional calendar.

Dosage data were also collected for all participants and are reported in Table 4.14. These data represent the number of learning trials presented to and completed by each participant within sessions. Dosage data provide a better understanding of participants' active role in their learning than can be provided by the amount of time spent in sessions. As is evident in Table 4.14, dosage varied considerably among participants and suggests that sessions were not equally productive.

Table 4.14

Dosage Summary Statistics by Participant

Participant	Min.	*Med*	*Mn*	*SD*	Max
1	14	20	24.93	7.19	40
2	10	20	24.62	8.77	40
3	20	100	90.46	34.43	200
4	20	80	72.91	24.37	100

Statistics and Data Analysis

Daily probe data for all participants are presented in Figures 4.1, 4.2, 4.3, and 4.4. Visual inspection of the data was conducted for both level and the immediacy of the change in performance after introduction of the intervention. With regard to the data's level, changes were observed for all participants between the baseline and intervention phases of the study. Baseline and intervention summary statistics for each participant are also reported in Table 4.15. CTOPP-2 post-testing data are reported in Table 4.16 while DIBELS-8 post-testing data and risk classifications are reported in Tables 4.17 and 4.18. Because participants one and two did not complete enough study sessions to have three opportunities to demonstrate an effect, visual analyses were not conducted with their data.

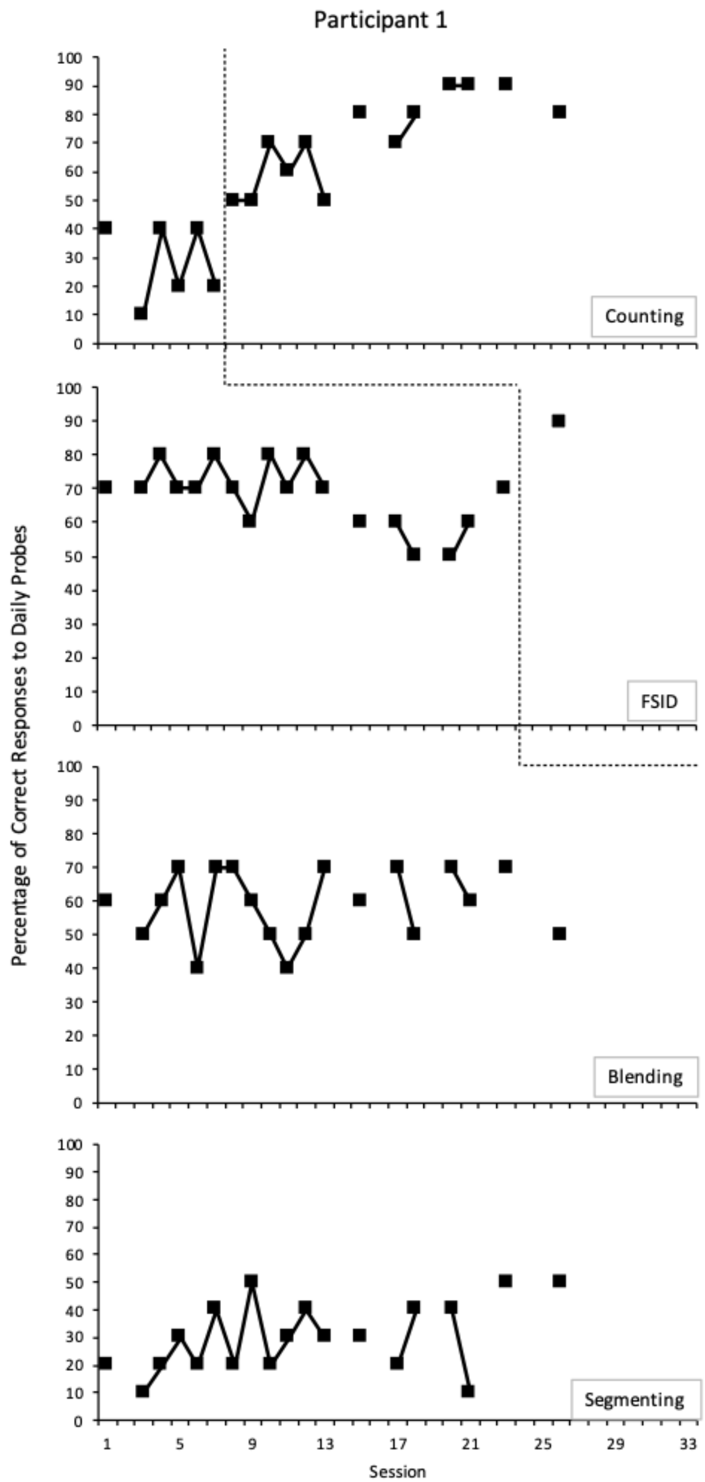

Figure 4.1. Probe accuracy percentages by session for participant one. *Note*. FSID refers to first phoneme identification.

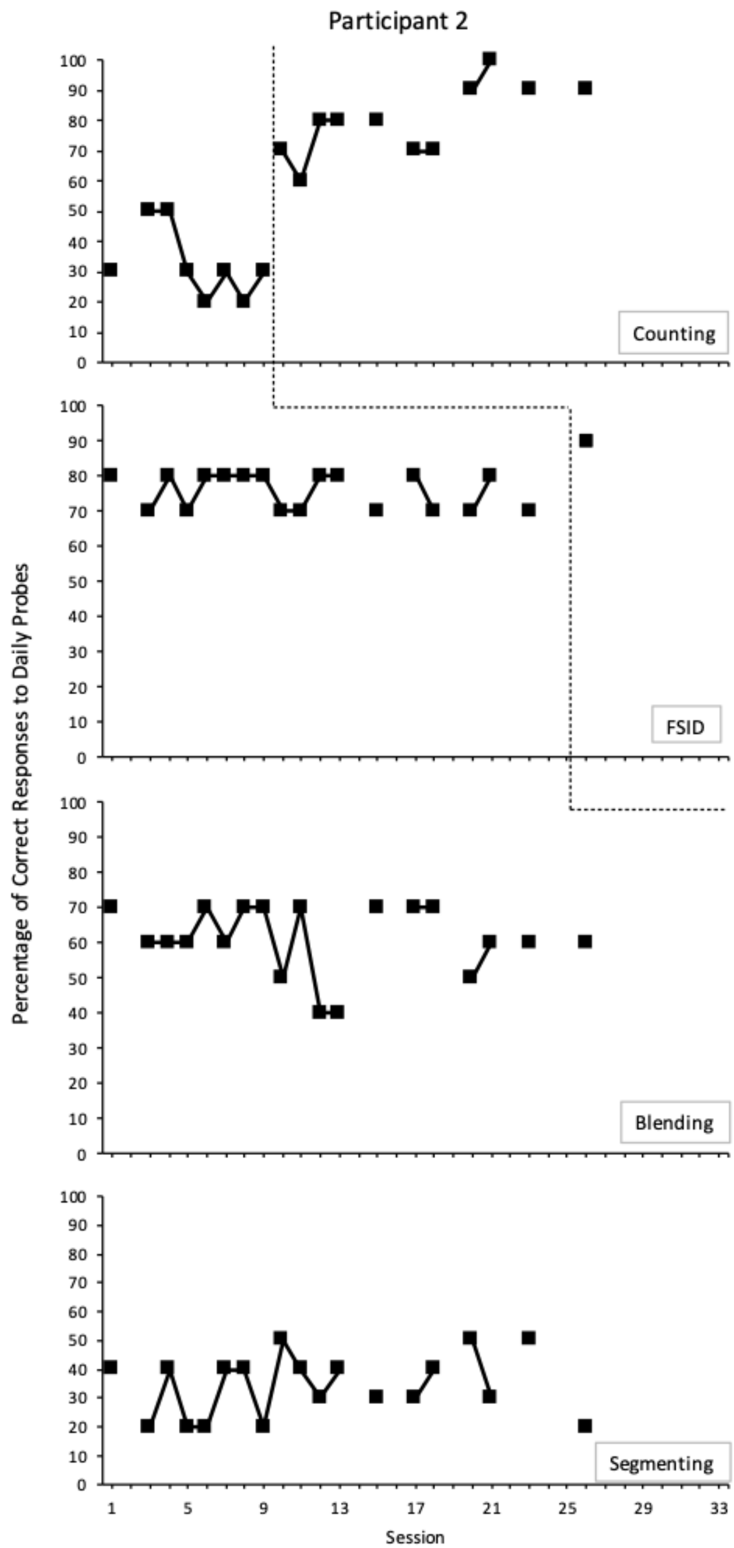

Figure 4.2. Probe accuracy percentages by session for participant two. *Note*. FSID refers to first phoneme identification.

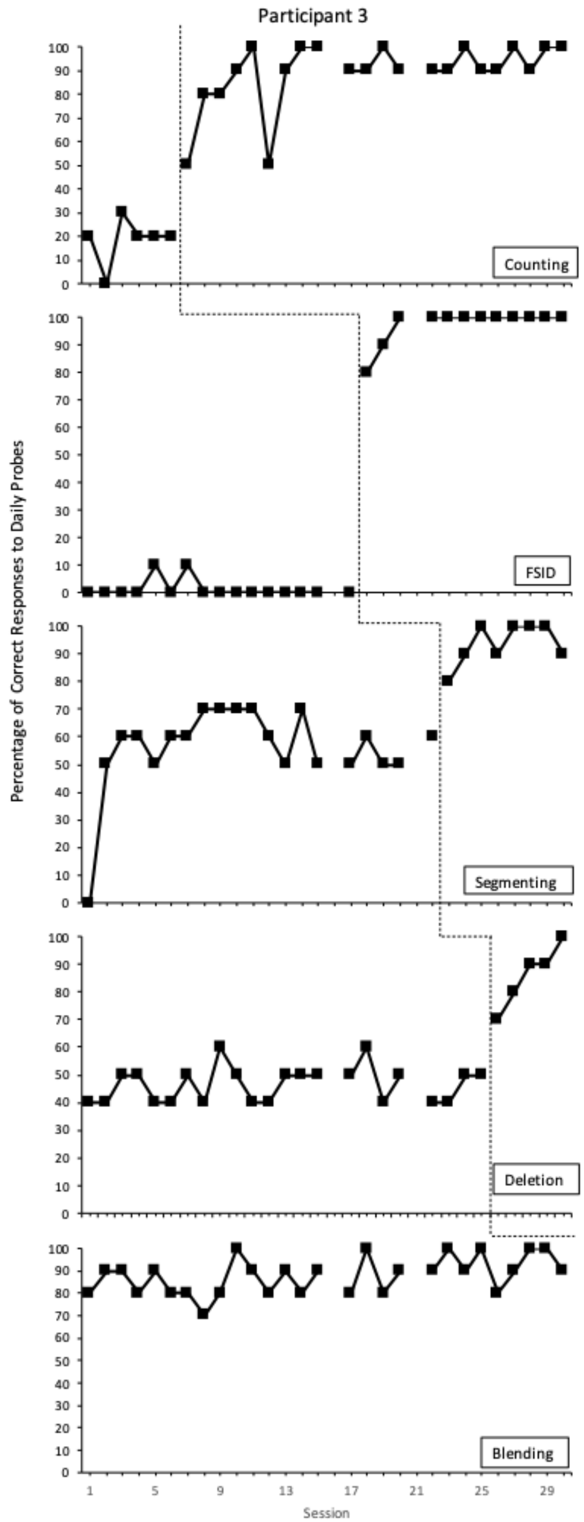

Figure 4.3. Probe accuracy percentages by session for participant three. *Note*. FSID refers to first phoneme identification.

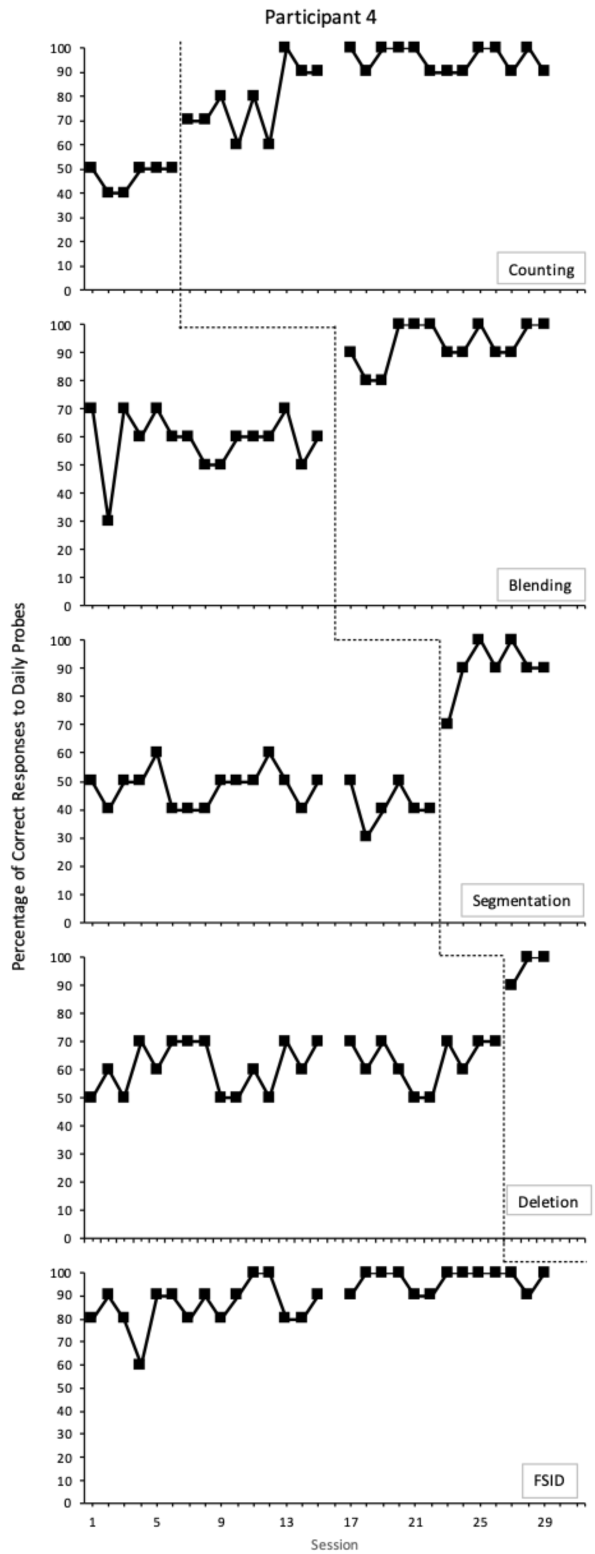

Figure 4.4. Probe accuracy percentages by session for participant four. *Note*. FSID refers to first phoneme identification.

Table 4.15

Baseline and Intervention Probe Data by Participant

	Baseline					Intervention				
Skill	Min.	*Mdn*	*Mn*	*SD*	Max	Min.	*Mdn*	*Mn*	*SD*	Max.
1										
Counting	10	30	28.33	13.29	40	50	70	71.54	15.19	90
FSID	50	70	67.78	9.43	80	90	90	90	—	90
Blending	40	60	58.95	10.49	70	—	—	—	—	—
Segmenting	10	30	30	12.91	50	—	—	—	—	—
2										
Counting	20	30	32.50	11.65	50	60	80	80	11.83	100
FSID	70	80	75.56	5.11	80	90	90	90	—	90
Blending	40	60	61.05	9.94	70	—	—	—	—	—
Segmenting	20	40	34.21	10.71	50	—	—	—	—	—
3										
Counting	0	20	18.33	9.83	30	50	90	89.09	14.11	100
FSID	0	0	1.25	3.42	10	80	100	97.50	6.22	100
Blending	70	90	87.86	8.33	100	—	—	—	—	—
Segmenting	0	60	56	15.36	70	80	95	93.75	7.44	100
Deletion	40	50	46.52	6.47	60	70	90	86	11.40	100
4										
Counting	40	50	46.67	5.16	50	60	90	88.18	12.96	100
FSID	60	90	90.71	9.79	100	—	—	—	—	—
Blending	30	60	58.67	10.60	70	80	90	93.08	7.51	100
Segmenting	30	50	46.19	7.40	60	70	90	90	10	100
Deletion	50	60	61.60	8.50	70	90	100	96.67	5.77	100

Table 4.16

CTOPP-2 Posttest Scores by Participant

Measure	Participant 3			Participant 4		
	Pretest	Posttest	Gain	Pretest	Posttest	Gain
Subtests						
Elision	6	12	6	7	9	2
Blending Words	8	14	6	9	14	5
Phoneme Isolation	5	12	7	7	13	6
Memory for Digits	5	5	0	5	5	0
Nonword Repetition	8	8	0	5	9	4
Rapid Digit Naming	9	8	-1	10	11	1
Rapid Letter Naming	7	8	1	10	9	-1
Blending Nonwords	7	13	6	8	13	5
Segmenting Nonwords	9	13	4	7	11	4
Composites						
Phonological Awareness	77	118	41	86	114	28
Phonological Memory	79	79	0	70	82	12
Rapid Symbolic Naming	88	88	0	101	101	0
Alt. Phonological Awareness	88	119	31	85	113	28

Table 4.17

DIBELS-8 Pretest, Posttest, and Gain Scores by Participant

Measure	Participant 3			Participant 4		
	Pretest	Posttest	Gain	Pretest	Posttest	Gain
PSF	46	68	22	55	73	18
NWF-CLS	59	78	19	40	56	16
NWF-WRC	12	25	13	8	16	8
ORF-ACC	88.88	97.53	8.65	78.57	91.30	12.73
ORF-WC	48	73	25	22	34	12
WRF	28.2	37	8.8	12	12	0
Maze Adjusted	2	11.5	9.5	2	5.5	3.5
Composite[a]	331.63	391.33	59.70	311.36	362.41	51.05

[a] The composite score is calculated using weighted formulas provided by the University of Oregon (2020) and does not include participants' PSF performance.

Table 4.18

DIBELS-8 Posttest Risk Classifications by Participant

Measure	Participant 3	Participant 4
PSF[a]	Negligible Risk	Negligible Risk
NWF-CLS	Min. Risk	Some Risk
NWF-WRC	Min. Risk*	Some Risk*
ORF-ACC	Min. Risk*	Some Risk*
ORF-WC	Some Risk	At Risk
WRF	Min. Risk	At Risk
Maze Adjusted	Min. Risk*	At Risk
Composite	Min. Risk	At Risk

[a] PSF risk classifications are reported based on the recommendations for children at the end of their first grade year.
* Asterisks represent a change in risk classification from pre-test.

Daily probe data. Immediate changes in level were noted for all participants on skills that were measured during both baseline and intervention phases with no overlap observed via visual analysis or a review of the summary statistics reported in Table 4.15. Statistical evaluation of baseline trends is reported in Table 4.20 in accordance with the methods used to calculate Tau (Tarlow, 2017), where *p*-values provide evidence regarding the rejection of the null hypothesis that a participant' baseline performance is stable. Across participants and skills reported here, baseline trends were identified by both visual and statistical analyses to be stable except as indicated in Table 4.20.

The daily probe data for participant three are presented graphically in Figure 4.3 and summary statistics for his performance are reported in Table 4.15. Across the baseline phases, performance on daily probes was low to medium for all skills (*Mdns*: 0, 20, 50, 60) with the exception of blending. Blending was subsequently not targeted for intervention due to high levels of accuracy demonstrated during baseline as indicated by a mean of 87.86% and standard deviation of 8.33%. For the remaining four skills, there was an increase in level seen immediately following the onset of the intervention. The participant met the mastery criterion for all four of the targeted skills. These data indicate there were four demonstrations of effect and zero non-demonstrations, supporting a functional relation between the implementation of the intervention and increases in his phonemic awareness skills as measured by daily probes.

The daily probe data for participant four are presented graphically in Figure 4.4 and summary statistics for his performance are also reported in Table 4.15. Across the baseline phases, performance on daily probes was in the medium range of ability for all skills (*Mdns*: 50, 50, 60, 60) with the exception of FSID. FSID was subsequently not targeted for intervention due to high levels of accuracy demonstrated during baseline as indicated by a mean of 90.71% and standard deviation of 9.79. For the remaining four skills, there was an increase in level seen immediately following the onset of the intervention. The participant met the mastery criterion for all four of the targeted skills. These data indicate there were four demonstrations of effect and zero non-demonstrations, supporting a functional relation between the implementation of the intervention and increases in his phonemic awareness skills as measured by daily probes.

Table 4.19

Effect Estimates by Participant

	Participant							
	1		2		3		4	
Skill	*SMD*	τ	SMD	τ	SMD	τ	SMD	τ
Counting	2.74	.709	3.62	.756	6.06	.667	6.77	.650
FSID	—	—	—	—	26.75	.893	—	—
Blending	—	—	—	—	—	—	3.07	.778
Segmenting	—	—	—	—	2.36	.710	5.70	.704
Deleting	—	—	—	—	5.89	.710	3.99	.518

Table 4.20

Baseline Trend Statistics by Participant

	Participant										
	1			2		3			4		
Skill	Tau	p		Tau	p	Tau	p		Tau	p	
Counting	-.078	1.000		-.423	.224	.086	1.000		.365	.487	
FSID	-.413	.032	*	-.199	.351	-.173	.475		.523	.001	*
Blending	.074	.713		-.126	.522	.346	.023	*	-.146	.522	
Segmenting	.388	.033	*	.140	.463	.025	.918		-.189	.303	
Deleting	—	—		—	—	.163	.359		.158	.339	

* Asterisks represent statistical significance at the .05 level.

Statistical analyses. Estimates of the standardized mean difference (SMD) and Tau were calculated using *R* using the SingleCaseES package (Pustejovsky & Swan, 2019) and code developed by Tarlow (2017), respectively. Tarlow's (2017) procedures were used for the calculation of Tau due to concerns regarding the limitations of procedures reported by Parker et al. (2011). Specifically, when using Parker et al.'s procedures to calculate Tau-U, values greater than one were obtained that made interpreting the true effect of the intervention difficult. As is indicated in Table 4.20., baseline correction procedures were not necessary for any of the skills with effects reported in Table 4.19 as none demonstrated statistically significant trends during the baseline phase of the study. The intervention was not introduced to any of the skills with statistically significant baseline trends as reported in Table 4.20. For Tau, benchmarks of .20, .60, and .80 have been offered as representing small, moderate, and large gains respectively (Vannest & Ninci, 2015). Together, the SMD and Tau values obtained across participants and skills indicate that contextualized phonemic awareness intervention resulted in a moderate to large improvement in skills over approximately seven weeks of individual intervention.

Post-test data. Gains were also observed on the CTOPP-2 and DIBELS-8 for both participants who completed the study as indicated in Tables 4.16-4.18. The current experimental design, however, does not allow for the conclusion of a functional relation between these gains and the intervention. Due to this lack of a functional relation, formal statistical analyses were not conducted on these measures and gains are reported above for descriptive and informational purposes only.

Discussion

The present study investigated the use of a contextualized phonemic awareness intervention program for children at risk of reading failure. In reviewing the results for both participants three and four, the present study suggests that contextualized phonemic awareness interventions can be highly effective at remediating phonemic awareness deficits during a critical period in the general education curriculum. Medium to large gains were observed in almost every skill targeted during the intervention with the exception of participant four's improvement in phoneme deletion representing only a small effect ($\tau = .518$). At the end of the study, both participants who completed the study demonstrated improvements in their their risk classification in three or four areas benchmarked by the DIBELS-8, although the current experimental design precludes . At the end of the study, participant four was still identified as "at risk" in four of six areas that were identified at the beginning of the study. He did, however, make improvements on his phoneme segmentation fluency, nonsense words read correctly, and the accuracy of his oral reading. His reading rate, as indicated by the ORF-WRC measure, was still slower than would be expected for a child of his age and suggests that he is still having to allocate substantial cognitive resources to decoding words. Although the number of words he read correctly in sixty seconds was still below expected levels, his improvement in *accurately* reading those words should not be discounted.

Both of the participants who completed the intervention were enrolled in the second grade at the time of this intervention, where the curriculum shifts from learning to read to reading to learn. Students who lack foundational reading skills, including phonemic awareness, are unlikely to successfully make this transition and will experience widening gaps between their own reading abilities and those that are expected of them as they move forward in formal schooling (Suggate, 2016). Interventions such as the contextualized phonemic awareness intervention investigated in the present study can be helpful in developing the requisite skills children need to be successful readers. As children develop their phonemic awareness and basic reading skills, contextualized interventions provide a natural segue between the development of phonemic awareness to the acquisition of basic word decoding and related recognition skills. Within contextualized interventions, children are able to contact reinforcement related to the experience of reading while receiving the necessary support to successfully practice their target skills.

As the present results suggest, work on phonemic awareness skills—even when contextualized—is not a panacea for children at risk of reading failure. Although improvements in multiple areas were seen in both participants who completed the intervention, both still needed continued work on reading fluency and participant four still demonstrated significant deficits in other areas as well. Although we did not formally assess meaningful engagement with the text, previous research suggests that contextualized phonemic awareness interventions can be helpful in connecting reading and pre-reading skill work to this end goal (e.g., Ukrainetz, 2015). Anecdotally, participant four in particular was observed to increase his attempts to decode and take a more active role in the joint reading process at the end of the intervention. During the last week of the intervention, he specifically asked to be allowed to attempt to read the text on his own with the principal investigator providing support as needed.

Overall, the results of the present study are congruent with the existing literature indicating the efficacy of phonemic awareness interventions (Ehri et al., 2001; Suggate, 2016). The present results extend the existing literature to include at-risk children who are older than have been previously included in the experimental literature under constraints common in applied settings. The observed effect sizes are also consistent with Suggate's (2016) findings that interventions for children in second grade should generally incorporate some degree of combined phonemic awareness and phonics instruction. By contextualizing the phonemic awareness tasks presented to participants using the written text of the books, participants' phonics knowledge was indirectly supported during each session of the present study. Furthermore, the observed effects are consistent with other studies in the literature indicating that phonemic awareness interventions implemented by speech-language pathologists can result in substantial gains (e.g., Gillon, 2000; Ukrainetz et al., 2009).

The present study indicates that speech-language pathologists can make substantial gains with at-risk children in a short period of time—in this case, approximately one quarter of a single school year. These results support a greater, more direct role for speech-language pathologists in RTI procedures for children at risk of reading failure. Due to school-based speech-language pathologists' ever-increasing workloads, however, future research should investigate how the effects of the present study translate when provided in a small group setting (Brandel, 2020). Future research should investigate how the effects of the current intervention are affected by its implementation in small or large group settings. Future research should also replicate the present study with more children and intervention sites, potentially utilizing a group research design based on the results of individual single case investigations.

Service Delivery Methods

Participants one and two participated in regularly scheduled sessions using a tablet and wireless hotspot provided to them by their school but connection issues were consistently a problem during sessions. Their guardian was present during most of their regularly scheduled sessions but maintaining their attention to session tasks remained difficult due to the free availability of competing activities in their home. Internet connectivity issues combined with off-task behavior substantially reduced the productivity of their sessions relative to participants three and four, as is evident above in Table 4.10. When comparing the productivity of their sessions to those of participants three and four, who received instruction via face-to-face instruction, it appears that there are more complex issues with the delivery of this intervention via distance learning methods.

Due to the myriad issues experienced when attempting to deliver instruction to participants one and two via teleconferencing software, future research should continue to investigate ways to mitigate the negative influence of these factors on student outcomes.

For example, although connectivity issues might be a continuing issue, better communication with caregivers at the beginning of and throughout interventions might help to resolve some of the issues related to scheduling and the availability of competing activities. It is also possible that researchers and clinicians might need to negotiate the intensity of intervention with caregivers to avoid placing too many demands on families' time if a lower intensity can provide comparable results.

Limitations

Although they missed over a third of their planned sessions, the administrative removal of participants one and two from the study is a limitation to the present work.

Participants three and four demonstrated higher moderately high performance on several skills at baseline, though not at the level of fluency representing mastery in the present work. Despite the lack of overlap between the intervention and baseline phases, this moderately high level of baseline performance can also be considered a limitation to the present study as the strength of the intervention is unclear. Future research should include participants with lower baseline levels of phonemic awareness to measure the strength of the intervention for children not on the cusp of mastery prior to intervention.

Finally, although data were collected on participants' generalized phonological awareness using the CTOPP-2 as well as their basic reading skills using the DIBELS-8, the relationship between the intervention and these outcomes is unclear given the current multiple baseline across behaviors research design. Future research should utilize a different experimental design that can more explicitly identify what relationship, if any, exists between contextualized phonemic awareness interventions and the development of basic literacy skills, specifically word decoding. Future research should likely also more directly incorporate graphemes into phonemic awareness training as is evidenced by the existing literature (e.g., Ehri et al., 2001; Suggate, 2016).

CHAPTER FIVE

Conclusion

The present work confirms that phonemic awareness instruction can be incredibly effective for children at risk of reading failure while simultaneously indicating that it is not equally effective in all situations. With previous research indicating that blending and segmenting outcomes are of the greatest importance relative to reading outcomes, the present meta-analysis helps to clarify the factors specifically affecting phonemic awareness by outcome measure. Specifically, small to medium outcomes were observed across all five outcomes investigated which is more consistent with the meta-analytic work of Suggate (2016) than Ehri et al. (2001). Although the sample sizes for subgroup analyses were small, this information can still be helpful for both researchers and practitioners seeking to better understand how to maximize the strength of phonemic awareness instruction when working with at-risk children.

The results of the meta-analysis suggest that children at-risk of reading failure can make large gains even when phonemic awareness instruction provided later than is generally recommended, though these effects do appear to decrease overall as children age. This is critical knowledge for researchers and practitioners working with children whose schools rely on whole language approaches to reading instruction that minimizes code-based instruction (Chaney, 1990). Schools also have considerable flexibility in the personnel they ask to administer these interventions, with adequate gains demonstrated when instruction is delivered by computers, paraprofessionals, community-based volunteers, teachers, and speech-language pathologists. With such a wide variety of providers available and capable of affecting positive change in students' phonemic awareness skills, schools can allocate their individual resources appropriately throughout the RTI and special education frameworks to best meet their community's learning needs. Consistent with that flexibility is the continued finding that individual, small group, and large group instructional arrangements are all capable of facilitating meaningful gains with this population.

Unsurprisingly, the single case investigation of contextualized phonemic awareness instruction agrees with the broad results of the meta-analysis: phonemic awareness instruction can affect significant change in at-risk children. Contextualizing phonemic awareness instruction into the activity of reading is consistent with the findings of Ehri et al. (2001), Suggate (2016), and the present meta-analytic review regarding the incorporation of graphemes with children who have developed some knowledge of sound-symbol correspondence. Because word reading is the goal of phonemic awareness instruction, incorporating written words into phonemic awareness instruction is a logical and meaningful bridge between the abstract task of developing phonemic awareness and the more practical work of decoding unknown words. The results of the present single case investigation, although it did not include a formal component of phonics instruction, suggest that even indirectly targeting phonics through contextualized phonemic awareness training might facilitate gains on basic reading skills other than just the phonemic awareness skills that were trained. Future research should more directly investigate this relationship using a different experimental design.

The gains observed in the present single case investigation are also consistent with the meta-analysis' findings that more intervention is not always better. In approximately seven weeks, both participants who completed the intervention made substantial gains in their phonemic awareness and other basic reading skills. Such substantial gains in such a short amount of time enables students to move forward towards interventions more exclusively targeting engagement with written text—that is, reading and writing words rather than orally manipulating them. As the American education system continues to grapple with the effects of school closures and instructional breaks related to covid-19 during the 2020 and 2021 school years, interventions that show the potential to affect significant change in basic literacy skills in a short amount of time will be critical components of the RTI framework.

Future research should build on the present work by investigating how the effects of the present single case intervention are affected when implemented in small or large group settings, when delivered by educators other than speech-language pathologists, or when implemented with children who have more severe phonemic awareness deficits. Future research could also compare the indirect incorporation of graphemes into phonemic awareness instruction used in the present study to a more direct incorporation of phonics instruction. The present work in conjunction with the research conducted over the past twenty years confirms that phonemic awareness instruction is highly effective with a wide variety of children, including those at risk of reading failure. Continuing to investigate variables related to how such interventions are designed and implemented is critical to the development of more efficient methods for the prevention and remediation of reading disorders in children.

As indicated in the present single case investigation, session productivity is not consistent across children and contexts. Future researchers need to ensure that they are reporting not only the length, frequency, and duration of intervention, but also relevant information regarding the dosage of instruction being delivered to each child within individual sessions. Such metrics are critical to understanding why interventions work for some children and not for others, as is evidenced by the discrepancy in outcomes between participants in the present study. As the current meta-analysis suggests that even small samples can demonstrate reliable effects of phonemic awareness interventions, researchers need to take care to report information that allows consumers to understand relevant variables impacting the effects of intervention with specific children. Dosage reporting may help to further clarify questions that remain after the present meta-analysis, such as the optimal number of skills to be taught within one intervention and the optimal session length, frequency, and duration of intervention for children at risk of reading failure. To date, information regarding the dosage of phonemic awareness instruction delivered to research participants is almost never reported for consumers to review, but should be readily available to researchers if they utilize data collection procedures similar to those used in the present single case investigation.

In summary, the present meta-analysis and single case investigation of phonemic awareness instruction provided to children at risk of reading failure confirm that this instruction is effective in helping these children acquire skills necessary for successful reading. Although phonemic awareness instruction is generally provided in kindergarten and first grade, the present work indicates that it can be effective with older children who lack these skills as well. Developing phonemic awareness in a contextualized manner such as in the present single case investigation, where the principal investigator maintained responsibility for reading the text and setting the foundation for the participants to engage in multiple phonemic awareness skills, can allow even older children to contact reinforcement while engaged with otherwise inaccessible texts.

APPENDIX

APPENDIX

List of Studies Included in Meta-Analysis

Adnams, C. M., Sorour, P., Kalberg, W. O., Kodituwakku, P., Perold, M. D., Kotze, A., September, S., Castle, B., Gossage, J., & May, P. A. (2007). Language and literacy outcomes from a pilot intervention study for children with fetal alcohol spectrum disorders in South Africa. *Alcohol*, *41*(6), 403-414. https://doi.org/10.1016/j.alcohol.2007.07.005

Al Otaiba, S., Schatschneider, C., & Silverman, E. (2005). Tutor-assisted intensive learning strategies in kindergarten: How much is enough? *Exceptionality*, *13*(4), 195-208. https://doi.org/10.1207/s15327035ex1304_2

Allor, J., & McCathren, R. (2004). The efficacy of an early literacy tutoring program implemented by college students. *Learning Disabilities Research and Practice*, *19*(2), 116–129. https://doi.org/10.1111/j.1540-5826.2004.00095.x

Anderson, J. A. (2009). *Parent-delivered instruction of phonemic awareness and letter identification skills for four- and five-year-old children in South Dakota*. University of South Dakota.

Balbi, A., & Cuadro, A. (2020). Efficacy of an evidence-based literacy intervention for Spanish speaking struggling readers from vulnerable socio-economic status. *Electronic Journal of Research in Educational Psychology*, *18*, 201–222. http://dx.doi.org/10.25115/ejrep.v18i50.2692

Barker, T. A. (1993). *An evaluation of computer-assisted instruction in phonological awareness with below average readers*. Florida State University.

Baugh, N. (2005). *The effect of risk level and group size on student phonemic awareness achievement*. Auburn University.

Benner, G. J. (2003). *An investigation of the effects of an intensive early literacy support program on the phonological processing skills of kindergarten children at-risk of emotional and behavioral disorders*. University of Nebraska.

Biancone, P. L. (2018). *The effects of an integrated early literacy and motor skill intervention on children's alphabet knowledge, initial sound awareness, and fundamental motor skill outcomes: An early efficacy study*. Ohio State University.

Bingham, G. E., Hall-Kenyon, K. M., & Culatta, B. (2010). Systematic and engaging early literacy: Examining the effects of paraeducator implemented early literacy instruction. *Communication Disorders Quarterly*, *32*(1), 38–49. https://doi.org/10.1177/1525740109340796

Björn, P. M., & Leppänen, P. H. T. (2013). Accelerating decoding-related skills in poor readers learning a foreign language: A computer-based intervention. *Educational Psychology*, *33*(6), 671–689. https://doi.org/10.1080/01443410.2013.797336

Bodé, S., & Content, A. (2011). Phonological awareness in kindergarten: A field study in Luxembourgish schools. *European Journal of Psychology of Education*, *26*(1), 109–128. https://doi.org/10.1007/s10212-010-0039-0

Bowyer-Crane, C., Snowling, M. J., Duff, F. J., Fieldsend, E., Carroll, J. M., Miles, J., Götz, K., & Hulme, C. (2008). Improving early language and literacy skills: Differential effects of an oral language versus a phonology with reading intervention. *Journal of Child Psychology and Psychiatry*, *49*(4), 422–432. https://doi.org/10.1111/j.1469-7610.2007.01849.x

Brady, S., Fowler, A., Stone, B., & Winbury, N. (1994). Training phonological awareness: A study with inner-city kindergarten children. *Annals of Dyslexia*, *44*(1), 26–59. https://doi.org/10.1007/BF02648154

Calhoon, M. B., Otaiba, S. A., Cihak, D., King, A., & Avalos, A. (2007). Effects of a peer-mediated program on reading skill acquisition for two-way bilingual first-grade classrooms. *Learning Disability Quarterly*, *30*(3), 169–184. https://doi.org/10.2307/30035562

Carson, K. L. (2020). Can an app a day keep illiteracy away? Piloting the efficacy of Reading Doctor apps for preschoolers with developmental language disorder. *International Journal of Speech-Language Pathology*, *22*(4), 454–465. https://doi.org/10.1080/17549507.2019.1667438

Carson, K. L., Bayetto, A. E., & Roberts, A. F. B. (2019). Effectiveness of preschool-wide teacher-implemented phoneme awareness and letter-sound knowledge instruction on code-based school-entry reading readiness. *Communication Disorders Quarterly*, *41*(1), 42–53. https://doi.org/10.1177/1525740118789061

Chera, P., & Wood, C. (2003). Animated multimedia 'talking books' can promote phonological awareness in children beginning to read. *Learning and Instruction*, *13*(1), 33–52. https://doi.org/10.1016/S0959-4752(01)00035-4

Compton, D. L. (2000). Modeling the response of normally achieving and at-risk first grade children to word reading instruction. *Annals of Dyslexia*, *50*(1), 53–84. https://doi.org/10.1007/s11881-000-0017-3

Crespo, P., Jiménez, J. E., Rodríguez, C., Baker, D., & Park, Y. (2018). Differences in growth reading patterns for at-risk Spanish-monolingual children as a function of a tier 2 intervention. *The Spanish Journal of Psychology*, *21*, 1–16. https://doi.org/10.1017/sjp.2018.3

DeBaryshe, B. D., & Gorecki, D. M. (2007). An experimental validation of a preschool emergent literacy curriculum. *Early Education & Development*, *18*(1), 93–110. https://doi.org/10.1080/10409280701274741

Duff, F. J., Hayiou-Thomas, M. E., & Hulme, C. (2012). Evaluating the effectiveness of a phonologically based reading intervention for struggling readers with varying language profiles. *Reading and Writing*, *25*(3), 621–640. https://doi.org/10.1007/s11145-010-9291-6

Ehri, L. C., Dreyer, L. G., Flugman, B., & Gross, A. (2007). Reading Rescue: An effective tutoring intervention model for language-minority students who are struggling readers in first grade. *American Educational Research Journal*, *44*(2), 414–448. https://doi.org/10.3102/0002831207302175

Elbro, C., & Petersen, D. K. (2004). Long-term effects of phoneme awareness and letter sound training: An intervention study with children at risk for dyslexia. *Journal of Educational Psychology*, *96*(4), 660–670. https://doi.org/10.1037/0022-0663.96.4.660

Epstein, J. N. (1994). *Accelerating the literacy development of disadvantaged preschool children: An experimental evaluation of a Head Start emergent literacy curriculum*. State University of New York at Stony Brook.

Fälth, L., Gustafson, S., & Svensson, I. (2017). Phonological awareness training with articulation promotes early reading development. *Education*, *137*, 261–267.

Fives, A., Kearns, N., Devaney, C., Canavan, J., Russell, D., Lyons, R., Eaton, P., & O'Brien, A. (2013). A one-to-one programme for at-risk readers delivered by older adult volunteers. *Review of Education*, *1*(3), 254–280. https://doi.org/10.1002/rev3.3016

Flis, A. (2006). *ABRACADABRA: The effectiveness of a computer-based reading intervention program for students at-risk for reading failure and students learning English as a second language*. University of Toronto.

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology*, *90*, 37–55. https://doi.org/10.1037/0022-0663.90.1.37

Foy, J. G. (2009). An effectiveness study of the Bookworm Project: A university-school partnership providing a low-cost, research-based early reading intervention for at-risk kindergarteners. *Early Childhood Services*, *3*, 281–300.

Fricke, S., Bowyer-Crane, C., Haley, A. J., Hulme, C., & Snowling, M. J. (2013). Efficacy of language intervention in the early years. *Journal of Child Psychology and Psychiatry*, *54*(3), 280–290. https://doi.org/10.1111/jcpp.12010

Fukuda, C. (2013). *Evaluation of a phonological awareness intervention for kindergarten and first grade students*. University of California Riverside.

Gillam, R. B., Loeb, D. F., Hoffman, L. M., Bohman, T., Champlin, C. A., Thibodeau, L., Widen, J., Brandel, J., & Friel-Patti, S. (2008). The efficacy of Fast ForWord Language intervention in school-age children with language impairment: A randomized controlled trial. *Journal of Speech, Language, and Hearing Research*, *51*(1), 97–119. https://doi.org/10.1044/1092-4388(2008/007)

Gillon, G. T. (2000). The efficacy of phonological awareness intervention for children with spoken language impairment. *Language, Speech, and Hearing Services in Schools*, *31*(2), 126–141. https://doi.org/10.1044/0161-1461.3102.126

Gillon, G. T. (2005). Facilitating phoneme awareness development in 3- and 4-year-old children with speech impairment. *Language, Speech, and Hearing Services in Schools*, *36*(4), 308–324. https://doi.org/10.1044/0161-1461(2005/031)

Gillon, G. T., Moran, C. A., Hamilton, E., Zens, N., Bayne, G., & Smith, D. (2007). Phonological awareness treatment effects for children from low socioeconomic backgrounds. *Asia Pacific Journal of Speech, Language and Hearing*, *10*(2), 123-140. https://doi.org/10.1179/136132807805297576

Given, B. K., Wasserman, J. D., Chari, S. A., Beattie, K., & Eden, G. F. (2008). A randomized, controlled study of computer-based intervention in middle school struggling readers. *Brain and Language*, *106*(2), 83–97. https://doi.org/10.1016/j.bandl.2007.12.001

Gonzalez, M. del R. O., Garcia Espinel, A. I., & Guzman Rosquete, R. (2002). Remedial interventions for children with reading disabilities: Speech perception—An effective component in phonological training? *Journal of Learning Disabilities*, *35*(4), 334–342. https://doi.org/10.1177/00222194020350040401

Hadley, P. A., Simmerman, A., Long, M., & Luna, M. (2000). Facilitating language development for inner-city children: Experimental evaluation of a collaborative, classroom-based intervention. *Language, Speech, and Hearing Services in Schools*, *31*(3), 280–295. https://doi.org/10.1044/0161-1461.3103.280

Hagans, K. S. (2013). Decreasing reading differences in children from disadvantaged backgrounds: The effects of an early literacy intervention. *Contemporary School Psychology*, *17*(1), 103–117.

Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development*, *65*, 41–57. https://doi.org/10.2307/1131364

Hatcher, P. J., Hulme, C., Miles, J. N. V., Carroll, J. M., Hatcher, J., Gibbs, S., Smith, G., Bowyer-Crane, C., & Snowling, M. J. (2005). Efficacy of small group reading intervention for beginning readers with reading-delay: A randomised controlled trial: Efficacy of small group reading intervention. *Journal of Child Psychology and Psychiatry*, *47*(8), 820–827. https://doi.org/10.1111/j.1469-7610.2005.01559.x

Hatcher, P. J., Hulme, C., & Snowling, M. J. (2004). Explicit phoneme training combined with phonic reading instruction helps young children at risk of reading failure. *Journal of Child Psychology and Psychiatry*, *45*(2), 338–358. https://doi.org/10.1111/j.1469-7610.2004.00225.x

Hecht, S. A., & Close, L. (2002). Emergent literacy skills and training time uniquely predict variability in responses to phonemic awareness training in disadvantaged kindergartners. *Journal of Experimental Child Psychology*, *82*(2), 93–115. https://doi.org/10.1016/S0022-0965(02)00001-2

Helf, S., Cooke, N. L., & Konrad, M. (2014). Advantages of providing structured supplemental reading instruction to kindergarteners at risk for failure in reading. *Preventing School Failure: Alternative Education for Children and Youth*, *58*(4), 214–222. https://doi.org/10.1080/1045988X.2013.798773

Hempenstall, K. (2008). Corrective reading: An evidence-based remedial reading intervention. *Australasian Journal of Special Education*, *32*(1), 23–54. https://doi.org/10.1080/10300110701839949

Hindson, B., Byrne, B., Fielding-Barnsley, R., Newman, C., Hine, D. W., & Shankweiler, D. (2005). Assessment and early instruction of preschool children at risk for reading disability. *Journal of Educational Psychology*, *97*(4), 687–704. https://doi.org/10.1037/0022-0663.97.4.687

Howell, R. D., Erickson, K., Stanger, C., & Wheaton, J. E. (2000). Evaluation of a computer-based program on the reading performance of first grade students with potential for reading failure. *Journal of Special Education Technology*, *15*(4), 5–14. https://doi.org/10.1177/016264340001500401

Innes, S. (1995). *Facilitation of phonological awareness in three and four year old children*. Queen Margaret College.

Iversen, S., & Tunmer, W. E. (1993). Phonological processing skills and the Reading Recovery Program. *Journal of Educational Psychology*, *85*(1), 112–126. https://doi.org/10.1037/0022-0663.85.1.112

Jimenez, J. E., del Rosario Ortiz, M., Rodrigo, M., Hernandez-Valle, I., Ramirez, G., Estevez, A., O'Shanahan, I., & de la Luz Trabaue, M. (2003). Do the effects of computer-assisted practice differ for children with reading disabilities with and without IQ-Achievement discrepancy? *Journal of Learning Disabilities*, *36*(1), 34–47. https://doi.org/10.1177/00222194030360010501

Kartal, G., Babür, N., & Erçetin, G. (2016). Training for phonological awareness in an orthographically transparent language in two different modalities. *Reading & Writing Quarterly*, *32*(6), 550–579. https://doi.org/10.1080/10573569.2015.1065213

Kartal, G., & Terziyan, T. (2016). Development and evaluation of game-like phonological awareness software for kindergarteners: JerenAli. *Journal of Educational Computing Research*, *53*(4), 519–539. https://doi.org/10.1177/0735633115608397

Kempert, S., Götz, R., Blatter, K., Tibken, C., Artelt, C., Schneider, W., & Stanat, P. (2016). Training early literacy related skills: To which degree does a musical training contribute to phonoogical awareness development? *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01803

Kerins, M. R., Trotter, D., & Schoenbrodt, L. (2010). Effects of a Tier 2 intervention on literacy measures: Lessons learned. *Child Language Teaching and Therapy*, *26*(3), 287–302. https://doi.org/10.1177/0265659009349985

Kleeck, A. van, Gillam, R. B., & McFadden, T. U. (1998). A study of classroom-based phonological awareness training for preschoolers with speech and/or language disorders. *American Journal of Speech-Language Pathology*, *7*(3), 65–76. https://doi.org/10.1044/1058-0360.0703.65

Korkman, M., & Peltomaa, A. K. (1993). Preventive treatment of dyslexia by a preschool training program for children with language impairments. *Journal of Clinical Child Psychology*, *22*, 277–287. https://doi.org/10.1207/s15374424jccp2202_13

Kyle, F., Kujala, J., Richardson, U., Lyytinen, H., & Goswami, U. (2013). Assessing the effectiveness of two theoretically motivated computer-assisted reading interventions in the United Kingdrom: GG Rime and GG Phoneme. *Reading Research Quarterly*, *48*(1), 61–76. https://doi.org/10.1002/rrq.038

Lane, H. B., Pullen, P. C., Hudson, R. F., & Konold, T. R. (2009). Identifying essential instructional components of literacy tutoring for struggling beginning readers. *Literacy Research and Instruction*, *48*(4), 277–297. https://doi.org/10.1080/19388070902875173

Lane, K. L., Fletcher, T., Carter, E. W., Dejud, C., & Delorenzo, J. (2007). Paraprofessional-led phonological awareness training with youngsters at risk for reading and behavioral concerns. *Remedial and Special Education*, *28*(5), 266–276. https://doi.org/10.1177/07419325070280050201

Layes, S., Guendouz, M., Lalonde, R., & Rebai, M. (2020). Combined phonological awareness and print knowledge training improves reading accuracy and comprehension in children with reading disabilities. *International Journal of Disability, Development and Education*, 1–15. https://doi.org/10.1080/1034912X.2020.1779914

Layne, D. D. (1997). *How does phonemic awareness affect reading ability with at-risk students on the primary level*. Rowan University.

Leafstedt, J. M., Richards, C. R., & Gerber, M. M. (2004). Effectiveness of explicit phonological-awareness instruction for at-risk English learners. *Learning Disabilities Research and Practice*, *19*(4), 252–261. https://doi.org/10.1111/j.1540-5826.2004.00110.x

Lennox, M., Westerveld, M. F., & Trembath, D. (2018). Evaluating the effectiveness of PrepSTART for promoting oral language and emergent literacy skills in disadvantaged preparatory students. *International Journal of Speech-Language Pathology*, *20*(2), 191–201. https://doi.org/10.1080/17549507.2016.1229030

Li, S., Tao, S., Joshi, R. M., & Xu, Q. (2018). Second-language reading difficulties among native Chinese-speaking students learning to read English: The roles of native- and second-language skills. *Reading Research Quarterly*, *53*(4), 423–441. https://doi.org/10.1002/rrq.204

Linan-Thompson, S., Pedrotty Bryant, D., Dickson, S. V., & Kouzekanani, K. (2005). Spanish literacy instruction for at-risk kindergarten students. *Remedial and Special Education*, *26*(4), 236–244. https://doi.org/10.1177/07419325050260040601

Lo, Y., Chuang Wang, & Haskell, S. (2009). Examining the Impacts of Early Reading Intervention on the Growth Rates in Basic Literacy Skills of At-Risk Urban Kindergarteners. *The Journal of Special Education*, *43*(1), 12–28. https://doi.org/10.1177/0022466907313450

Loeb, D. F., Gillam, R. B., Hoffman, L., Brandel, J., & Marquis, J. (2009). The effects of Fast ForWord Language on the phonemic awareness and reading skills of school-age children with language impairments and poor reading skills. *American Journal of Speech-Language Pathology*, *18*(4), 376–387. https://doi.org/10.1044/1058-0360(2009/08-0067)

Lonigan, C. J., Driscoll, K., Phillips, B. M., Cantor, B. G., Anthony, J. L., & Goldstein, H. (2003). A computer-assisted instruction phonological sensitivity program for preschool children at-risk for reading problems. *Journal of Early Intervention*, *25*(4), 248–262. https://doi.org/10.1177/105381510302500402

Lonigan, C. J., & Phillips, B. M. (2016). Response to instruction in preschool: Results of two randomized studies with children at significant risk of reading difficulties. *Journal of Educational Psychology*, *108*(1), 114–129. https://doi.org/10.1037/edu0000054

Lonigan, C. J., Purpura, D. J., Wilson, S. B., Walker, P. M., & Clancy-Menchetti, J. (2013). Evaluating the components of an emergent literacy intervention for preschool children at risk for reading difficulties. *Journal of Experimental Child Psychology*, *114*(1), 111–130. https://doi.org/10.1016/j.jecp.2012.08.010

MacKay, L. D. (2007). *Assessing the effectiveness of a cultural curriculum to improve early literacy outcomes for kindergarteners*. University of British Columbia.

Makhoul, B. (2017). Moving beyond phonological awareness: The role of phonological awareness skills in Arabic reading development. *Journal of Psycholinguistic Research*, *46*(2), 469–480. https://doi.org/10.1007/s10936-016-9447-x

Mathes, P. G., Torgesen, J. K., & Allor, J. H. (2001). The effects of peer-assisted literacy strategies for first-grade readers with and without additional computer-assisted instruction in phonological awareness. *American Educational Research Journal*, *38*(2), 371–410. https://doi.org/10.3102/00028312038002371

McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children*, *71*(4), 445–463. https://doi.org/10.1177/001440290507100404

Milburn, T. F., Lonigan, C. J., & Phillips, B. M. (2017). Determining responsiveness to tier 2 intervention in response to intervention: Level of performance, growth, or both. *The Elementary School Journal*, *118*(2), 310–334. https://doi.org/10.1086/694271

Mirbazel, S. K., Arjmandi, M., & Bazargani, D. T. (2016). The effect of phoneme awareness of English sounds on the improvement of reading ability of Iranian junior secondary program (JSP) students. *Modern Journal of Language Teaching Methods*, *6*(4), 186–198.

Mitchell, M. J., & Fox, B. J. (2001). The effects of computer software for developing phonological awareness in low-progress readers. *Reading Research and Instruction*, *40*(4), 315–332. https://doi.org/10.1080/19388070109558353

Nancollis, A., Lawrie, B.-A., & Dodd, B. (2005). Phonological awareness intervention and the acquisition of literacy skills in children from deprived social backgrounds. *Language, Speech, and Hearing Services in Schools*, *36*(4), 325–335. https://doi.org/10.1044/0161-1461(2005/032)

Nutkins, C. (2004). *Examining the effectiveness of phonemic awareness training for at-risk preschool children: Does parental involvement promote skill development?* Central Michigan University.

O'Connor, M., Arnott, W., McIntosh, B., & Dodd, B. (2009). Phonological awareness and language intervention in preschoolers from low socio-economic backgrounds: A longitudinal investigation. *British Journal of Developmental Psychology*, *27*(4), 767–782. https://doi.org/10.1348/026151008X372492

O'Connor, R. E., Bocian, K., Beebe-Frankenberger, M., & Linklater, D. L. (2010). Responsiveness of students with language difficulties to early intervention in reading. *The Journal of Special Education*, *43*(4), 220–235. https://doi.org/10.1177/0022466908317789

Olofsson, A., & Lundberg, I. (1985). Evaluation of long term effects of phonemic awareness training in kindergarten: Illustrations of some methodological problems in evaluation research. *Scandinavian Journal of Psychology*, *26*, 21–34. https://doi.org/10.1111/j.1467-9450.1985.tb01138.x

Olson, R. K., Wise, B., Ring, J., & Johnson, M. (1997). Computer-based remedial training in phoneme awareness and phonological decoding: Effects on the posttraining development of word recognition. *Scientific Studies of Reading*, *1*(3), 235–253. https://doi.org/10.1207/s1532799xssr0103_4

O'Shaugnessy, T. (1997). *A comparison of two reading interventions for children with or at-risk for reading disabilities*. University of California Riverside.

Pietrangelo, D. (1999). *Outcomes of an enhanced literacy curriculum on the emergent literacy skills of Head Start preschoolers*. University at Albany, State University of New York.

Pindiprolu, S. S., & Marks, L. J. (2020). Preventing summer reading slide: Examining the effects of two computer-assisted reading programs. *Rural Special Education Quarterly*, *39*(3), 116–127. https://doi.org/10.1177/8756870520914281

Plasencia-Peinado, J. (1998). *Examining the effect of three Spanish reading approaches on the reading achievement of first-grade limited English proficient (LEP) students at-risk for special education*. University of Oregon.

Pokorni, J. L., Worthington, C. K., & Jamison, P. J. (2004). Phonological awareness intervention: Comparison of Fast ForWord, Earobics, and LiPS. *The Journal of Educational Research*, *97*(3), 147–158. https://doi.org/10.3200/JOER.97.3.147-158

Poskiparta, E., Niemi, P., & Vauras, M. (1999). Who benefits from training in linguistic awareness in the first grade, and what components show training effects? *Journal of Learning Disabilities*, *32*(5), 437–446. https://doi.org/10.1177/002221949903200509

Pullen, P. C., & Lane, H. B. (2014). Teacher-directed decoding practice with manipulative letters and word reading skill development of struggling first grade students. *Exceptionality*, *22*(1), 1–16. https://doi.org/10.1080/09362835.2014.865952

Raisor, L. J. (2006). *A comparison of phonological awareness intervention approaches*. University of Cincinatti.

Regtvoort, A. G. F. M., & van der Leij, A. (2007). Early intervention with children of dyslexic parents: Effects of computer-based reading instruction at home on literacy acquisition. *Learning and Individual Differences*, *17*(1), 35–53. https://doi.org/10.1016/j.lindif.2007.01.005

Ritter, M. J., Park, J., Saxon, T. F., & Colson, K. A. (2013). A phonologically based intervention for school-age children with language impairment: Implications for reading achievement. *Journal of Literacy Research*, *45*(4), 356–385. https://doi.org/10.1177/1086296X13504156

Ron Nelson, J., Benner, G. J., & Gonzalez, J. (2005). An investigation of the effects of a prereading intervention on the early literacy skills of children at risk of emotional disturbance and reading problems. *Journal of Emotional and Behavioral Disorders*, *13*(1), 3–12. https://doi.org/10.1177/10634266050130010101

Rule, A. C., Dockstader, C. J., & Stewart, R. A. (2006). Hands-on and kinesthetic activities for teaching phonological awareness. *Early Childhood Education Journal*, *34*(3), 195–201. https://doi.org/10.1007/s10643-006-0130-y

Ryder, J. F., Tunmer, W. E., & Greaney, K. T. (2008). Explicit instruction in phonemic awareness and phonemically based decoding skills as an intervention strategy for struggling readers in whole language classrooms. *Reading and Writing*, *21*(4), 349–369. https://doi.org/10.1007/s11145-007-9080-z

Scanlon, D. M., Gelzheiser, L. M., Vellutino, F. R., Schatschneider, C., & Sweeney, J. M. (2008). Reducing the incidence of early reading difficulties: Professional Development for classroom teachers versus direct interventions for children. *Learning and Individual Differences*, *18*(3), 346–359. https://doi.org/10.1016/j.lindif.2008.05.002

Schneider, W., Ennemoser, M., Roth, E., & Küspert, P. (1999). Kindergarten prevention of dyslexia: Does training in phonological awareness work for everybody? *Journal of Learning Disabilities*, *32*(5), 429–436. https://doi.org/10.1177/002221949903200508

Schneider, W., Roth, E., & Ennemoser, M. (2000). Training phonological skills and letter knowledge in children at risk for dyslexia: A comparison of three kindergarten intervention programs. *Journal of Educational Psychology*, *92*, 284–295. https://psycnet.apa.org/doi/10.1037/0022-0663.92.2.284

Schwartz, R. M. (2005). Literacy learning of at-risk first-grade students in the reading recovery early intervention. *Journal of Educational Psychology*, *97*(2), 257–267. https://doi.org/10.1037/0022-0663.97.2.257

Segers, E., & Verhoeven, L. (2004). Computer-supported phonological awareness intervention for kindergarten children with specific language impairment. *Language, Speech, and Hearing Services in Schools*, *35*(3), 229–239. https://doi.org/10.1044/0161-1461(2004/022)

Sénéchal, M., Ouellette, G., Pagan, S., & Lever, R. (2012). The role of invented spelling on learning to read in low-phoneme awareness kindergartners: A randomized-control-trial study. *Reading and Writing*, *25*(4), 917–934. https://doi.org/10.1007/s11145-011-9310-2

Seward, C., & Laurier, W. (2009). *Evaluating the effectiveness of a short-duration reading intervention on grade one phonological awareness and word reading*. Wilfred Laurier University.

Simmons, D. C., Kame'enui, E. J., Harn, B., Coyne, M. D., Stoolmiller, M., Edwards Santoro, L., Smith, S. B., Beck, C. T., & Kaufman, N. K. (2007). Attributes of effective and efficient kindergarten reading intervention: An examination of instructional time and design specificity. *Journal of Learning Disabilities*, *40*(4), 331–347. https://doi.org/10.1177/00222194070400040401

Solari, E. J., Denton, C. A., Petscher, Y., & Haring, C. (2018). Examining the Effects and Feasibility of a Teacher-Implemented Tier 1 and Tier 2 Intervention in Word Reading, Fluency, and Comprehension. *Journal of Research on Educational Effectiveness*, *11*(2), 163–191. https://doi.org/10.1080/19345747.2017.1375582

Solari, E. J., & Gerber, M. M. (2008). Early comprehension instruction for Spanish-speaking English language learners: Teaching text-level reading skills while maintaining effects on word-level skills. *Learning Disabilities Research & Practice*, *23*(4), 155–168. https://doi.org/10.1111/j.1540-5826.2008.00273.x

Stevens, R. J., Van Meter, P. N., Garner, J., Warcholak, N., Bochna, C., & Hall, T. (2008). Reading and Integrated Literacy Strategies (RAILS): An integrated approach to early reading. *Journal of Education for Students Placed at Risk (JESPAR)*, *13*(4), 357–380. https://doi.org/10.1080/10824660802427611

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1997). Prevention and remediation of severe reading disabilities: Keeping the end in mind. *Scientific Studies of Reading*, *1*, 217–234. https://doi.org/10.1207/s1532799xssr0103_3

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Herron, J., & Lindamood, P. (2010). Computer-assisted instruction to prevent early reading difficulties in students at risk for dyslexia: Outcomes from two instructional approaches. *Annals of Dyslexia*, *60*(1), 40–56. https://doi.org/10.1007/s11881-009-0032-y

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., & Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology*, *91*(4), 579–593. https://doi.org/10.1037/0022-0663.91.4.579

Trainin, G., Wilson, K. M., Murphy-Yagil, M., & Rankin-Erickson, J. L. (2014). Taking a different route: Contribution of articulation and metacognition to intervention with at-risk third-grade readers. *Journal of Education for Students Placed at Risk (JESPAR)*, *19*(3–4), 183–195. https://doi.org/10.1080/10824669.2014.972103

Travis, P. (1997). *Effects of computer-assisted and teacher-led phonological awareness instruction for first-grade students at risk for reading failure*. University of Florida.

Troia, G. A., & Whitney, S. D. (2003). A close look at the efficacy of Fast ForWord Language for children with academic weaknesses. *Contemporary Educational Psychology*, *28*(4), 465–494. https://doi.org/10.1016/S0361-476X(02)00045-0

Ukrainetz, T. A., Cooney, M. H., Dyer, S. K., Kysar, A. J., & Harris, T. J. (2000). An investigation into teaching phonemic awareness through shared reading and writing. *Early Childhood Research Quarterly*, *15*(3), 331–355. https://doi.org/10.1016/S0885-2006(00)00070-3

Ukrainetz, T. A., Ross, C. L., & Harm, H. M. (2009). An investigation of treatment scheduling for phonemic awareness with kindergartners who are at risk for reading difficulties. *Language, Speech, and Hearing Services in Schools*, *40*(1), 86–100. https://doi.org/10.1044/0161-1461(2008/07-0077)

Vadasy, P. F., Jenkins, J. R., Antil, L. R., Wayne, S. K., & O'Connor, R. E. (1997). The effectiveness of one-to-one tutoring by community tutors for at-risk beginning readers. *Learning Disability Quarterly*, *20*(2), 126–139. https://doi.org/10.2307/1511219

Vadasy, P. F., Jenkins, J. R., & Pool, K. (2000). Effects of tutoring in phonological and early reading skills on students at risk for reading disabilities. *Journal of Learning Disabilities*, *33*(6), 579–590. https://doi.org/10.1177/002221940003300606

Vadasy, P. F., Sanders, E. A., & Peyton, J. A. (2006). Code-oriented instruction for kindergarten students at risk for reading difficulties: A randomized field trial with paraeducator implementers. *Journal of Educational Psychology*, *98*(3), 508–528. https://doi.org/10.1037/0022-0663.98.3.508

van Otterloo, S. G., & van der Leij, A. (2009). Dutch home-based pre-reading intervention with children at familial risk of dyslexia. *Annals of Dyslexia*, *59*(2), 169–195. https://doi.org/10.1007/s11881-009-0030-0

Vanderwood, M. L., Tung, C., & Arellano, E. (2014). Application of a multitiered system of support with English language learners. *International Journal of School & Educational Psychology*, *2*(1), 45–53. https://doi.org/10.1080/21683603.2013.860570

Vaughn, S., Cirino, P. T., Linan-Thompson, S., Mathes, P. G., Carlson, C. D., Hagan, E. C., Pollard-Durodola, S. D., Fletcher, J. M., & Francis, D. J. (2006). Effectiveness of a Spanish intervention and an English intervention for English-language learners at risk for reading problems. *American Educational Research Journal*, *43*(3), 449–487. https://doi.org/10.3102/00028312043003449

Vaughn, S., Mathes, P., Linan-Thompson, S., Cirino, P., Carlson, C., Pollard-Durodola, S., Cardenas-Hagan, E., & Francis, D. (2006). Effectiveness of an English intervention for first-grade English language learners at risk for reading problems. *The Elementary School Journal*, *107*(2), 153–180. https://doi.org/10.1086/510653

Wang, C.-C. (1999). *Learning to read Chinese: The roles of phonological awareness and morphological awareness*. University of Illinoi at Urbana-Champaign.

Warren, P. (2009). *The effects of training parents in teaching phonemic awareness on the phonemic awareness and early reading of struggling readers*. Auburn University.

Warrick, N., Rubin, H., & Rowe-Walsh, S. (1993). Phoneme awareness in language-delayed children: Comparative studies and intervention. *Annals of Dyslexia*, *43*(1), 153–173. https://doi.org/10.1007/BF02928179

Weiner, S. (1994). Effects of phonemic awareness training on low- and middle-achieving first graders' phonemic awareness and reading ability. *Journal of Reading Behavior*, *26*(3), 277–300. https://doi.org/10.1080/10862969409547852

Wheldall, R., Glenn, K., Arakelian, S., Madelaine, A., Reynolds, M., & Wheldall, K. (2016). Efficacy of an evidence-based literacy preparation program for young children beginning school. *Australian Journal of Learning Difficulties*, *21*(1), 21–39. https://doi.org/10.1080/19404158.2016.1189443

Whittaker, S. (2013). *Response to Intervention (RtI) effectiveness in kindergarten reading achievement*. Aurora University.

Wise, B. W., Ring, J., & Olson, R. K. (1999). Training phonological awareness with and without explicit attention to articulation. *Journal of Experimental Child Psychology*, *72*(4), 271–304. https://doi.org/10.1006/jecp.1999.2490

Wise, B. W., Ring, J., & Olson, R. K. (2000). Individual differences in gains from computer-assisted remedial reading. *Journal of Experimental Child Psychology*, *77*(3), 197–235. https://doi.org/10.1006/jecp.1999.2559

Wise, N. (2014). *Phonological awareness training for struggling readers in grade 1 French immersion*. University of Toronto.

Wise, N., D'Angelo, N., & Chen, X. (2016). A school-based phonological awareness intervention for struggling readers in early French immersion. *Reading and Writing*, *29*(2), 183–205. https://doi.org/10.1007/s11145-015-9585-9

Wolff, U. (2011). Effects of a randomised reading intervention study: An application of structural equation modeling. *Dyslexia*, *17*(4), 295–311. https://doi.org/10.1002/dys.438

Wolff, U. (2014). RAN as a predictor of reading skills, and vice versa: Results from a randomised reading intervention. *Annals of Dyslexia*, *64*(2), 151–165. https://doi.org/10.1007/s11881-014-0091-6

Yeh, S. S. (2003). An evaluation of two approaches for teaching phonemic awareness to children in Head Start. *Early Childhood Research Quarterly*, *18*(4), 513–529. https://doi.org/10.1016/j.ecresq.2003.09.009

Yeh, S. S., & Connell, D. B. (2008). Effects of rhyming, vocabulary and phonemic awareness instruction on phoneme awareness. *Journal of Research in Reading*, *31*(2), 243–256. https://doi.org/10.1111/j.1467-9817.2007.00353.x

Zoski, J. L. (2015). *Kindergarten linguistic awareness intervention: The effects of phonological, morphological, and orthographic instruction on early literacy outcomes in at-risk students*. University of North Carolina at Chapel Hill.

REFERENCES

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

Adnams, C. M., Sorour, P., Kalberg, W. O., Kodituwakku, P., Perold, M. D., Kotze, A., . . . May, P. A. (2007). Language and literacy outcomes from a pilot intervention study for children with fetal alcohol spectrum disorders in South Africa. *Alcohol, 41*, 403-414.

Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention: A review of the literature. *Remedial and Special Education, 23*, 300-316.

Al Otaiba, S., Schatschneider, C., & Silverman, E. (2005). Tutor-assisted intensive learning strategies in kindergarten: How much is enough? *Exceptionality*, 195-208.

Alonzo, C. N., McIlraith, A. L., Catts, H. W., & Hogan, T. P. (2020). Predicting dyslexia in children with developmental language disorder. *Journal of Speech, Language, and Hearing Research, 63*, 151-162.

Anderson, J. A. (2010). *Parent-delivered instruction of phonemic awareness and letter identification skills for four- and five-year-old children in South Dakota [Unpublished dissertation]*. University of South Dakota.

Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence-based Mental Health, 22*, 153-160.

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50*, 1088-1101.

Björn, P. M., & Leppänen, P. H. (2013). Accelerating decoding-related skills in poor readers learning a foreign language: A computer-based intervention. *An International Journal of Experimental Educational Psychology, 33*, 671-689.

Brandel, J. (2020). Speech-language pathology services in the schools: A follow-up 9 years later. *Language, Speech, and Hearing Services in Schools, 51*, 1037-1048.

Braun, E. (2012). *Trust me, Jack's beanstalk stinks! The story of Jack and the beanstalk as told by the giant.* Picture Window Books.

Buehner, C. (2002). *Snowmen at Night.* Dial Books for Young Readers.

Bus, A. G., Ijzendoorn, M. H., & Pellegrini, A. (1995). Joint book reading makes success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research, 65*, 1-21.

Caffrey, E., Fuchs, D., & Fuchs, L. S. (2008). The predictive validity of dynamic assessment. *The Journal of Special Education, 41*, 254-270.

Card, N. A. (2012). *Applied meta analysis for social science research.* Guilford.

Carson, K. L. (2020). Can an app a day keep illiteracy away? Piloting the efficacy of Reading Doctor apps for preschoolers with developmental language disorder. *International Journal of Speech-Language Pathology, 22*, 454-465.

Carson, K. L., Bayetto, A. E., & Roberts, A. F. (2019). Effectiveness of preschool-wide teacher-implemented phoneme awareness and letter-sound knowledge instruction on code-based school-entry reading readiness. *Communication Disorders Quarterly, 41*, 42-53.

Catts, H. W. (2009). The case for the narrow view of reading. *Language, Speech, and Hearing Services in Schools, 40*, 174-177.

Catts, H. W., Alof, S. M., Hogan, T. P., & Weismer, S. E. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language, and Hearing Research, 48*, 1378-1396.

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, Z. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research, 45*, 1142-1157.

Chaney, C. (1990). Evaluating the whole language approach to language arts. *Language, Speech, and Hearing Services in Schools, 21*, 244-249.

Chera, P., & Wood, C. (2003). Animated multimedia 'talking books' can promote phonological awareness in children beginning to read. *Learning and Instruction, 13*, 33-52.

Cochrane. (2021). *Cochrane handbook for systematic reviews of interventions (Version 6.2).* (J. P. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch, Eds.) Retrieved from https://training.cochrane.org/handbook

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.).* Lawrence Erlbaum Associates.

Eagle, J., Dowd-Eagle, S. E., Snyder, A., & Holtzman, E. G. (2015). Implementing a multi-tiered system of support (MTSS): Collaboration between school psychologists and administrators to promote systems-level change. *Journal of Educational and Psychological Consultation, 25*, 160-177.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*, 629-634.

Ehri, L. C. (2002). Phases of acquisition in learning to read words and implications for teaching. In R. Stainthorp, & P. Tomlinson, *Learning and teaching reading* (pp. 7-28). British Journal of Educational Psychology.

Ehri, L. C., Nunes, S. R., Willows, D. M., Valeska Schuster, B., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36*(250-287).

Field, A. (2018). *Discovering statistics using IBM SPSS Statistics (5th ed.).* SAGE.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Stuebing, K. K., . . . Shaywitz, B. A. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology, 86*, 6-23.

Flis, A. (2006). *ABRACADABRA: The effectiveness of a computer-based reading intervention program for students at-risk for reading failure, and students learning English as a second language [Unpublished dissertation].* University of Toronto.

Frizelle, P., Tolonen, A., Tulip, J., Murphy, C., Saldana, D., & McKean, C. (2021). The influence of quantitative intervention dosage on oral language outcomes for children with developmental language disorder: A Systematic review and narrative synthesis. *Language, Speech, and Hearing Services in Schools, 52*, 768-754.

Fuchs, L. S., & Fuchs, D. (2002). Principles for the prevention and intervention of mathematics difficulties. *Learning Disabilities Research and Practice, 16*, 85-95.

Gillam, R. B., McFadden, T. A., & van Kleeck, A. (1995). Improving the narrative abilities of children with language disorders: Whole languae and language skills approaches. In M. Fey, J. Windsor, & S. Warren, *Childhood language disorders* (pp. 145-182). Paul H. Brookes.

Gillam, S. L., Gillam, R. B., Reece, K., Nippold, M., & Schneider, P. (2012). Language outcomes of contextualized and decontextualized language intervention: Results of an early efficacy study. *Language, Speech, and Hearing Services in Schools, 43*, 276-291.

Gillon, G. T. (2000). The efficacy of phonological awareness intervention for children with spoken language impairment. *Language, Speech, and Hearing Services in Schools, 31*, 126-141.

Given, B. K., Wasserman, J. D., Chari, S. A., Beattie, K., & Eden, G. F. (2008). A randomized, controlled study of computer-based intervention in middle school struggling readers. *Brain and Language, 106*, 83-97.

Given, B. K., Wasserman, J. D., Chari, S. A., Beattie, K., & Eden, G. F. (2008). A randomized, controlled stuy of computer-based intervention in middle school struggling readers. *Brain and Language, 106*, 83-97.

Goldsworthy, C. L. (2012). *The sourcebook of phonological awareness activities (2nd ed.)* (Vol. 1). Delmar.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*, 6-10.

Gunderson, J. (2016). *Frankly, I'd rather spin myself a new name! The story of Rumpelstiltskin as told by Rumpelstiltskin.* Picture Window Books.

Gunderson, J. (2016). *No lie, pigs (and their houses) can fly! The story of the three little pigs as told by the wolf.* Picture Window Books.

Hall, M. S., & Burns, M. K. (2018). Meta-analysis of targeted small-group reading interventions. *Journal of School Psychology, 66*, 54-66.

Hartmann, D. P., Barrios, B. A., & Wood, D. D. (2004). Principles of behavioral observation. In S. N. Haynes, & E. M. Hieby, *Comprehensive handbook of psychological assessment.* Wiley.

Hatcher, P. J., Hulme, C., Miles, J. N., Carroll, J. M., Hatcher, J., Gibbs, S., . . . Snowling, M. J. (2006). Efficacy of small group reading intervention for beginning readers with reading-delay: A randomised controlled trial. *Journal of Child Psychology and Psychiatry, 47*, 820-827.

Hempenstall, K. (2008). Corrective reading: An evidence-based remedial reading intervention. *Australasian Journal of Special Education, 32*, 23-54.

Hindson, B., Byrne, B., Fielding-Barnsley, R., Newman, C., Hine, D. W., & Shankweiler, D. (2005). Assessment and early instruction of preschool children at risk for reading disability. *Journal of Educational Psychology, 97*, 687-704.

Hogan, T. P., Catts, H. W., & Little, T. D. (2005). The relationship between phonological awareness and reading: Implications for the assessment of phonological awareness. *Language, Speech, and Hearing Services in Schools, 36*, 285-293.

Howell, R. D., Erickson, K., Stanger, C., & Wheaton, J. E. (2000). Evaluation of a computer-based program on the reading performance of first grade students with potential for reading failure. *Journal of Special Education Technology, 15*, 5-14.

Justice, L. M. (2006). Evidence-based practice, response to intervention, and the prevention of reading difficulties. *Language, Speech, and Hearing Services in Schools, 37*, 284-297.

Justice, L. M., & Ezell, H. K. (2004). Print referencing: An emergent literacy enhancement strategy and its clinical applications. *Language, Speech, and Hearing Services in Schools, 35*, 185-193.

Kamhi, A. G. (2002). The path to proficient word recognition. *Perspectives on School-Based Issues, 3*, 43-47.

Kamhi, A. G., & Catts, H. L. (2012). *Language and reading disabilities (3rd ed.).* Pearson.

Kamhi, A. G., & Catts, H. W. (2017). Reading comprehension is not a single ability—Implications for assessment and instruction. *Language, Speech, and Hearing Servics in Schools, 48*, 104-107.

Kartal, G., Babür, N., & Erçetin, G. (2016). Training for phonological awareness in an orthographically transparent language in two different modalities. *Reading & Writing Quarterly, 32*, 550-579.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test (2nd ed.).* AGS Publishing.

Kazdin, A. E. (2011). *Single case research designs: Methods for clinical and applied settings (2nd ed.).* Oxford University Press.

Kerins, M. R., Trotter, D., & Schoenbrodt, L. (2010). Effects of a Tier 2 intervention on literacy measures: Lessons learned. *Child Language Teaching and Therapy, 26*, 287-302.

Kranzler, J. H., Gilbert, K., Robert, C. R., Floyd, R. G., Benson, N. F., & Reschly, A. (2019). Further examination of a critical assumption underlying the dual-discrepancy/consistency approach to specific learning disability identification. *School Psychology Review, 48*, 207-221.

Lane, K. L., Fletcher, T., Carter, E. W., Dejud, C., & DeLorenzo, J. (2007). Paraprofessional-led phonological awareness training with youngsters at risk for reading and behavioral concerns. *Remedial and Special Education, 28*, 266-276.

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

Lo, Y., Wang, C., & Haskell, S. (2009). Examining the impacts of early reading intervention on the growth rates in basic literacy skills of at-risk urban kindergarteners. *Journal of Special Education, 43*, 12-28.

Loewen, N. (2012). *Believe me, Goldilocks rocks! The story of the three bears as told by baby bear.* Picture Window Books.

Loewen, N. (2013). *Seriously, Snow White was so forgetful! The story of Snow white as told by the dwarves.* Picture Window Books.

Loewen, N. (2016). *Trust me, Hansel and Gretel are sweet! The story of Hansel and Gretel as told by the witch.* Picture Window Books.

Loewen, N. (2018). *Listen, my bridge is so cool! The story of the three billy goats gruff as told by the troll.* Picture Window Books.

McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children, 71*, 445-463.

McNeill, B. C., Justice, L. M., Gillon, G. T., & Schuele, C. M. (2017). Phonological awareness: Description, assessment, and intervention. In J. E. Bernthal, N. W. Bankson, & P. Flipsen, *Articulation and phonological disorders (8th ed.)* (pp. 302-324). Pearson.

Melby-Lervåg, M., Lyster, S. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin, 138*, 1092-1098.

Mirbazel, S. K., Arjmandi, M., & Bazargani, D. T. (2016). The effect of phonemic awareness of English sounds on the improvement of reading ability of Iranian junior secondary students. *Modern Journal of Language Teaching Methods, 6*, 186-198.

Nancollis, A., Lawrie, B., & Dodd, B. (2005). Phonological awareness intervention and the acquisition of literacy skills in children from deprived social backgrounds. *Language, Speech, and Hearing Services in Schools, 36*, 325-335.

Nelson, J. R., Benner, G. J., & Gonzalez, J. (2005). An investigation of the effects of a prereading intervention on the early literacy skills of children at risk of emotional disturbance and reading problems. *Journal of Emotional and Behavioral Disorders, 13*, 3-12.

Office of Special Education and Rehabilitative Services. (2011). *A Response to Intervention (RTI) process cannot be used to delay-deny an evaluation for eligibility under the Individuals with Disabilities Education Act (IDEA).*

O'Shaugnessy, T. E. (1997). *A comparison of two reading interventions for children with or at-risk for reading disabilities [Unpublished dissertation].* University of California Riverside.

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research. *Behavior Therapy, 42*, 284-299.

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

Peterson, D. B. (2011). A systematic review of narrative-based language intervention with children who have language impairment. *Communication Disorders Quarterly, 32*, 207-220.

Piasta, S. B., Justice, L. M., McGinty, A. S., & Kaderavek, J. N. (2012). Increasing young children's contact with print during shared reading: Longitudinal effects on literacy achievement. *Child Development, 83*, 810-820.

Pustejovsky, J. E., & Swan, D. M. (2019). *SingleCaseES: A Calculator for Single-Case Effect Sizes (R code)*. Retrieved from https://CRAN.R-project.org/package=SingleCaseES

Renaissance Learning. (2017). *STAR Early Literacy Assessment.*

Ryder, J. F., Tunmer, W. E., & Greaney, K. T. (2008). Explicit instruction in phonemic awareness and phonemically based decoding skills as an intervention strategy for struggling readers in whole language classrooms. *Reading and Writing, 21*, 349-369.

Schuele, M. C., & Boudreau, D. (2008). Phonological awareness intervention: Beyond the basics. *Language, Speech, and Hearing Services in Schools, 39*, 3-20.

Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). *Meta-analysis with R.* Springer.

Shaskan, T. S. (2012). *Honestly, Red Riding Hood was rotten! The story of Little Red Riding Hood as told by the wolf.* Picture Window Books.

Spector, J. E. (1995). Phonemic awareness training: Application of principles of direct instruction. *Reading & Writing Quarterly, 11*, 37-51.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 86*, 360-406.

Sterett, B. I., McDaniel, S. C., Majeika, C. E., & Bruhn, A. L. (2020). Using evidence informed strategies to adapt tier 2 interventions. *Journal of Applied School Psychology, 36*, 133-154.

Stevens, R. J., Van Meter, P. N., Garner, J., Warcholak, N., Bochna, C., & Hall, T. (2008). Reading and integrated litearcy strategies (RAILS): An integrated approach to early reading. *Journal of Education for Students Placed At Risk, 13*, 357-380.

Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities, 49*, 77-96.

Swanson, L. A., Fey, M. E., Mills, C. E., & Hood, L. S. (2005). Use of narrative-based language intervention with children who have specific language impairment. *American Journal of Speech-Language Pathology, 14*, 131-141.

Tarlow, K. R. (2017). An improved rank correlation effect size statistic for single-case designs: Baseline corrected tau. *Behavior Modification, 41*, 427-467.

Tarlow, K. R. (2017, June). *Baseline corrected Tau for single-case research (R code)*. Retrieved from http://ktarlow.com/stats/

Tiruchittampalam, S., Nicholson, T., Levin, J. R., & Ferron, J. M. (2018). The effects of preliteracy knowledge, schooling, and summer vacation on literacy acquisition. *Journal of Educational Research, 111*, 28-42.

Torgesen, J. K., Rashotte, C. A., Alexander, A., Alexander, J., & MacPhee, K. (2003). Progress towards understanding the instructional conditions necessary for remediating reading difficulties in older children. In B. Foorman, *Preventing and remediating reading difficulties: Bringing science to scale* (pp. 275-298). York Press.

Torgesen, J., Wagner, R., & Rashotte, C. (1994). Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities, 27*, 276-286.

Troia, G. A. (1999). Phonological awareness intervention research: A critical review of the experimental methodology. *Reading Research Quarterly, 34*, 28-52.

Troia, G. A., & Whitney, S. D. (2003). A close look at the efficacy of Fast ForWord Language for children with academic weaknesses. *Contemporary Educational Psychology, 28*, 465-494.

Ukrainetz, T. A. (2015). *School-age language intervention: Evidence based practices.* Pro-Ed.

Ukrainetz, T. A., Ross, C. L., & Harm, H. M. (2009). An investigation of treatment scheduling for phonemic awareness with kindergartners who are at risk for reading difficulties. *Language, Speech, and Hearing Services in Schools, 40*, 86-100.

University of Oregon. (2020). *DIBELS-8 administration and scoring guide.* Author.

Vanderwood, M. L., Tung, C., & Arellano, E. (2014). Application of a multitiered system of support with English language learners. *International Journal of School and Educational Psychology, 2*, 45-53.

Vannest, K. J., & Ninci, J. (2015). Evaluating intervention effects in single-case research designs. *Journal of Counseling and Development, 93*, 403-411.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*, 1-48.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing (2nd ed.).* Pro-Ed.

What Works Clearinghouse. (2020). *Standards handbook (Version 4.1).* Retrieved from https://ies.ed.gov/ncee/wwc/Handbooks

Wheldall, R., Glenn, K., Arakelian, S., Madelaine, A., Reynolds, M., & Wheldall, K. (2016). Efficacy of an evidence-based literacy preparation program for young children beginning school. *Australian Journal of Learning Difficulties, 21*, 21-39.

Wise, B. W., Ring, J., & Olson, R. K. (1999). Training phonological awareness with and without explicit attention to articulation. *Journal of Experimental Child Psychology, 72*, 271-304.

Wise, N. (2016). *Phonological awareness training for struggling readers in grade 1 French immersion [Unpublished dissertation].* University of Toronto.

Wolff, U. (2011). Effects of a randomised reading intervention study: An application of structural equation modeling. *Dyslexia, 17*, 295-311.

Wren, Y., Hambly, H., & Roulstone, S. (2013). A review of the impact of bilingualism on the development of phonemic awareness skills in children with typical speech development. *Child Language Teaching and Therapy, 29*, 11-25.

Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly, 23*, 159-177.