Published in final edited form as:

J Learn Disabil. 2018 ; 51(6): 612–624. doi:10.1177/0022219418775110.

Current Evidence on the Effects of Intensive Early Reading Interventions

Jeanne Wanzek, PhD[1], **Elizabeth A. Stevens, MAEd**[2], **Kelly J. Williams, MEd, EdS**[2], **Nancy Scammacca, PhD**[2], **Sharon Vaughn, PhD**[2], and **Katherine Sargent, MEd**[1]

[1]Vanderbilt University, Nashville, TN, USA

[2]The University of Texas at Austin, Austin, TX, USA

Abstract

Many students at risk for or identified with reading disabilities need intensive reading interventions. This meta-analysis provides an update to the Wanzek and Vaughn synthesis on intensive early reading interventions. Effects from 25 reading intervention studies are analyzed to examine the overall effect of intensive early reading interventions as well as relationships between intervention and student characteristics related to outcomes. The weighted mean effect size (ES) estimate (ES = 0.39), with a mean effect size adjusted for publication bias (ES = 0.28), both significantly different from zero, suggested intensive early reading interventions resulted in positive outcomes for early struggling readers in kindergarten through third grades. There was no statistically significant or meaningful heterogeneity in the study-wise effect sizes. Exploratory examination of time in intervention, instructional group size, initial reading achievement, and date of publication are provided.

Keywords

elementary; age; intervention; reading

While intensive interventions can be implemented in any academic domain, reading is a primary target for intensive interventions in the primary grades. The primary grades are particularly important because many reading difficulties and disabilities can be prevented if students are provided with early reading intervention (Fuchs, Compton, Fuchs, Bryant, & Davis, 2008; Partanen & Siegel, 2014; Simmons et al., 2008; Torgesen et al., 2001; Vellutino, Scanlon, Small, & Fanuele, 2006). One way to identify students in need of intensive intervention is to examine reading achievement prior to intervention. Research has noted that very low levels of initial reading achievement predict later low levels of reading achievement even when these students are provided less intensive Tier 2 type interventions (Al Otaiba & Fuchs, 2002; Lam & McMaster, 2014; Nelson, Benner, & Gonzalez, 2003;

Corresponding Author: Jeanne Wanzek, PhD, Department of Special Education, Vanderbilt University, 110 Magnolia Circle, Nashville, TN 37203, USA., jeanne.wanzek@vanderbilt.edu.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Vaughn & Linan-Thompson, 2003). Deficits in phonological awareness, rapid naming, fluency, and the alphabetic principle appear to be the most consistent predictors of initial response to intervention (Al Otaiba & Fuchs, 2002; Lam & McMaster, 2014; Nelson et al., 2003).

Furthermore, syntheses of the impact of reading interventions provided in the primary grades report higher average impacts on reading outcomes than interventions implemented in the upper elementary and secondary grades (Scammacca, Roberts, Vaughn, & Stuebing, 2015; Wanzek et al., 2013; Wanzek & Vaughn, 2007; Wanzek, Wexler, Vaughn, & Ciullo, 2010). If students do not develop strong reading skills in the primary grades, they will most likely continue to have difficulty reading throughout school (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Juel, 1988; McNamara, Scissons, & Gutknectch, 2011; Stanovich, 1986) and will be at an increased risk for dropping out of school (Hernandez, 2011).

Over the past 10 years, there has been an increased emphasis on providing intensive reading interventions. In 2011, the Office of Special Education Programs funded the National Center on Intensive Intervention (www.intensiveintervention.org), which provides technical assistance to districts to support their use of intensive interventions. Their work includes the use of technical review committees to evaluate intervention programs, assessment tools, implementation strategies, and the use of intensive interventions in current schools. Furthermore, the Division for Learning Disabilities (DLD) published a position statement recommending that students at risk for and with learning disabilities (LD) receive intensive evidence-based interventions in areas of need in addition to accommodations and modifications in order to be able to fully access the general curriculum (Vaughn, Zumeta, Wanzek, Cook, & Klingner, 2014).

One way previous research has conceptualized intensive intervention is to provide more time in intervention. Increasing session length or the duration of the intervention may allow students with significant reading needs to accelerate their learning (Denton, Fletcher, Anthony, & Francis, 2006; Torgesen et al., 2001; Vaughn, Linan-Thompson, & Hickman, 2003). Other studies have provided a more intensive reading intervention by reducing the size of the instructional group for intervention (Lou et al., 1996; Vaughn, Linan-Thompson, Kouzekanani, et al., 2003). In these studies, students receiving reading intervention in small groups outperformed students receiving reading intervention in larger groups. This research on intensifying interventions through time, dosage, or instructional group size became the basis for the quantitative ways to increase intensity of intervention that are commonly used in instructional models with multitiered intervention levels (National Center on Intensive Intervention, 2012). In 2008, the Institute of Education Sciences summarized recent reading intervention research and recommended that students with minimal response to Tier 2 interventions receive intensive daily reading intervention (Gersten et al., 2008). However, this recommendation had low evidence when evaluated using the What Works Clearinghouse standards. Only five studies were available for review at that time, and there were no significant effects, indicating a clear need for updated review of the response to intervention (RTI) literature, particularly in the area of intensive reading interventions. The Center on Instruction released a practice guide on intensive interventions in reading and

mathematics (Vaughn, Wanzek, Murray, & Roberts, 2012). This guide provided specific recommendations based on the available research for intensifying interventions through (a) the use of strategies that promote cognitive processes, (b) delivering more explicit and systematic instruction in addition to increased opportunities for feedback, (c) providing additional instructional time, and (d) decreasing group size (Vaughn et al., 2012). This guide emphasized the way in which intensive interventions could be implemented in school settings; however, it did not review the efficacy studies of intensive reading interventions.

Wanzek and Vaughn (2007) conducted a systematic review of intensive reading interventions (i.e., those with 100 sessions or more) in Grades K through 3 from 1995 to 2005 to inform the implementation of intensive interventions within an RTI framework. They identified 13 studies that met criteria for inclusion in the review and determined that the intensive interventions resulted in generally positive effects on reading outcomes (mean effect size [ES] range, –0.05 to 0.84). Additionally, the researchers synthesized intervention effects by duration, group size, grade level, and level of individualization, although these variables were not systematically manipulated in the original studies. The duration of the intervention did not appear to be related to the magnitude of the intervention effect. Studies of intensive interventions taught individually (1:1) reported greater effects than those taught in groups (two to eight students). There were not enough studies with small groups (two to four students) for the authors to examine the effects of small group implementation. Grade level was related to the magnitude of intervention effect, with larger effects reported in studies in Grades K through 1 than in Grades 2 and 3. Wanzek and Vaughn also sought to examine the effects of standardized and individualized interventions, but they were unable to locate any intensive reading interventions utilizing an individualized approach to instruction. There were no differences in effects on student reading outcomes for studies implementing highly standardized versus less standardized interventions.

Additional research on intensive reading interventions has been conducted since Wanzek and Vaughn (2007) was published that can provide updated information on the effects of these intensity variables in reading intervention.

Purpose for and Rationale of Study

Many students at risk for or identified with reading disabilities need intensive reading interventions (Fuchs, Fuchs, & Vaughn, 2014; Vaughn, Denton, & Fletcher, 2010; Vaughn & Wanzek, 2014). Practical recommendations for intensifying reading interventions continue to include examination of intervention duration, instructional group size, and individualization (National Center on Intensive Intervention, 2012). Since the Wanzek and Vaughn (2007) synthesis on intensive early reading interventions, additional research has been conducted examining the efficacy of reading interventions in Grades K through 3. The current synthesis aims to update and extend the findings from the previous synthesis on intensive early reading interventions (Wanzek & Vaughn, 2007). We systematically searched the literature for studies conducted since the end of the last synthesis search (2005) using the same criteria as Wanzek and Vaughn to provide an updated corpus of studies on intensive reading interventions from 1995 to 2015. We extended the previous synthesis by conducting a meta-analysis of the effects of intensive interventions and sought to examine duration of

intervention, instructional group size, grade level, and individualization of the intervention as moderators of these effects, just as were examined originally in Wanzek and Vaughn . We also included publication year as a moderator to examine possible change over time in the effects of the research from the original synthesis to the present. Finally, we extended the previous synthesis by examining the effects of intervention by initial reading achievement given its general relationship to intervention responsiveness in the research. Specifically, we addressed three research questions:

Research Question 1: What are the effects of intensive early reading interventions for students with reading difficulties?

Research Question 2: What intervention or student characteristics are related to student outcomes?

Research Question 3: Does publication date predict the magnitude of effects of intensive early reading interventions for students with reading difficulties?

Method

A comprehensive search of the literature was conducted to identify articles that met inclusion criteria. First, we included all of the identified studies from the initial synthesis (Wanzek & Vaughn, 2007) with sufficient data for calculating effect sizes. Second, we conducted an updated search of PsycINFO and ERIC to identify studies published since the original synthesis (2006 and 2015). Abstracts were searched utilizing terms related to reading intervention (*reading interven*, reading instruction, reading strategies, supplemental instruction, special educ*, phon*, fluency, vocab*, comp**) in combination with search terms intended to target our key population (*reading difficult*, learning disab*, reading disab*, reading delays, reading disorder*, dyslex**). Additionally, a hand search was conducted of relevant reading- and learning disabilities–related journals (*Exceptional Children, Journal of Learning Disabilities, Journal of Special Education, Learning Disabilities Research and Practice, Reading Research Quarterly, Scientific Studies of Reading, Reading and Writing, Journal of Research in Reading, Journal of Literacy Research, Reading Teacher, Reading and Writing Quarterly, Learning Disability Quarterly, Remedial and Special Education, Dyslexia, Annals of Dyslexia)* in 2015 to ensure all recently published articles were reviewed.

Our initial updated search identified 8,039 abstracts for screening. We applied the same inclusion criteria used in Wanzek and Vaughn (2007):

1. The study was reported in a peer-reviewed journal and printed in English.
2. The participants included students with learning disabilities or students identified as at risk for reading difficulties (e.g., students with low ability, low phonemic awareness, language disorders); studies with additional participants were included if disaggregated data were provided for the students with learning disabilities or the students were identified as at risk.
3. The participants were enrolled in Grades K through 3 inclusive.
4. Interventions targeted early literacy in an alphabetic language.

5. Instruction was provided as part of school programming (i.e., not including home, clinic, or camp programs).

6. Reading outcomes were measured in the study.

In addition to the criteria applied from the Wanzek and Vaughn (2007) synthesis, we also selected only those studies with sufficient data to calculate effect sizes between study groups so that we could conduct the meta-analysis.

Abstract review led to disqualification of 7,638 studies from the initial updated search (e.g., not intervention studies, age of participants outside of criterion). The full text of the remaining 401 articles was read. Of those articles, 13 studies met all inclusion criteria. These 13 studies from the updated search along with 12 studies from Wanzek and Vaughn (2007) with sufficient data to calculate effect sizes between study groups were included. Thus, a total of 25 studies were included in the meta-analysis.

Coding Procedures

A detailed coding sheet was utilized to organize relevant information about each study (Vaughn, Elbaum, Wanzek, Scammacca, & Walker, 2014). The code sheet was based on elements specified in the What Works Clearinghouse Design and Implementation Assessment Device (Institute of Education Sciences, 2008) as well as coding information utilized in previous research. Data were collected on (a) participants (e.g., socioeconomic status, risk type, gender), (b) research design and methodology, (c) description of experimental and control conditions, (d) clarity of causal inference, (e) measures, and (f) results. There were four trained coders. A gold standard method of coding was used with each coder demonstrating reliability to the first author before beginning coding. Intercoder agreement was calculated separately for each category on the code sheet (e.g., participants, measures, results) and reached 90% or above for all categories (range, 92%–100%). Intercoder agreement was calculated as the number of agreements divided by the number of agreements plus disagreements. In addition, all code sheets were independently coded by two coders. Any discrepancies were discussed until consensus was reached.

Effect Size Calculation

For all studies, Hedges's *g* was calculated using the means and standard deviations for each group when such data were provided. Cohen's *d* effect sizes and sample sizes for each group were used to calculate Hedges's *g* when means and standard deviations were not reported. All effects were computed using the Comprehensive Meta Analysis (Version 3.3.070) software (Borenstein, Hedges, Higgins, & Rothstein, 2013).

Meta-Analysis Procedures

Standardized versus unstandardized measures.—Based on previous research that showed that effect sizes in reading intervention studies differ in magnitude based on whether the outcome measures are standardized or not (Scammacca et al., 2015; Swanson, Hoskyn, & Lee, 1999), we considered conducting separate meta-analyses for each type of measure. However, in the corpus of studies that met the inclusion criteria, only 24 of the 328 effect sizes were calculated from scores on unstandardized measures. All of the 24 instances

involved studies where both standardized and unstandardized measures were used to estimate treatment effects. Therefore, we decided to include both types of measures in the meta-analysis and conduct a sensitivity analysis to determine if results would differ if the unstandardized measures were not included.

Treatment-comparison versus multiple treatment studies.—Additionally, we decided to meta-analyze effect sizes from both studies that used a treatment-comparison contrast ($k = 19$) and studies that contrasted two or more treatments ($k = 6$). This decision was based on several factors. First, the comparison condition in the treatment-comparison studies typically was business-as-usual classroom instruction, not a no-treatment control. Therefore, it seemed reasonable to view the comparison group's instruction as an alternate treatment (albeit one not dictated by the researchers). Also, when mean effect sizes for the two study designs were compared, no significant difference was found ($p = .76$), leading us to determine that including both types of studies likely would not add systematic heterogeneity to the meta-analytic results. Finally, including both types of studies provided additional power for the meta-analysis.

Analytic approach.—Given that 328 effect sizes were calculated from 25 studies, it is apparent that most studies contained data for computing more than one effect size. The mean number of effect sizes per study was 13.12, with a range of 3 to 42. Multiple effect sizes resulted from multiple outcome measures being used to estimate the treatment effect and from studies that involved more than one pair of group contrasts. When multiple effect sizes are computed from a single study, the meta-analytic data contain dependencies that must be accounted for in the analysis.

To accommodate the dependency in the data, robust variance estimation (RVE; Hedges, Tipton, & Johnson, 2010) was implemented to conduct the meta-analysis, using the *robumeta* package for R (Fisher & Tipton, 2015). RVE adjusts the standard errors of the effect size estimates to account for the correlation between effect sizes within studies. RVE requires that the mean correlation between all pairs of effect sizes within a study (ρ) be specified to estimate the study weights and calculate the between-study variance. As shown by Hedges et al. (2010), the value selected for ρ generally does not yield meaningful differences in the results; they recommended conducting a sensitivity analysis in which varying ρ values are used in the meta-analytic models. Using 0.2, 0.5, and 0.8, we found no differences in the results of the meta-analysis. The results reported in the following used a ρ of 0.8.

An additional consideration when using RVE is the increase in Type I error rates that has been found when the number of studies included in the meta-analysis is less than 40 (Tipton, 2015). Because this meta-analysis included 25 studies, the small-sample correction developed by Tipton (2015) was implemented when the model was run in *robumeta*. Tipton showed that this correction prevents inflation of Type I error rates. An intercept-only meta-regression model was run to estimate the overall mean effect size and standard error and calculate indices of heterogeneity.

Moderator analyses.—We planned to conduct moderator analyses using the following variables in meta-regression models implemented with RVE: hours of treatment (less than or equal to 63 hours vs. more than 63 hours), instructional group size (small group vs. one on one), grade level, individualization of the intervention, year of publication, and ability level of the sample based on pretest standardized reading test scores. However, as Wanzek and Vaughn (2007) also encountered, the majority of the work on intensive interventions occurred in Grades K through 1 and examined only standardized interventions, preventing examination of grade level and individualization of intervention as moderators of effects due to lack of variation across studies. Hours of treatment could not be operationalized as a continuous variable because this information was reported as a range or mean in many studies. As a result, we chose the median of the data set as a dividing point to categorize the length of treatment. One study had a wide range of total hours of instruction that crossed the median and was therefore not included in the analysis for total hours. Standardized pretest scores were reported in 12 of the 25 studies. For these studies, a *z* score was calculated using the mean and standard deviation of the test's normative sample to estimate the ability level of the study's sample. This *z* score was used as a covariate in a meta-regression model to determine if pretest ability level predicted the posttest effect size for the difference between groups. We recognized that power for the moderator analyses was low and considered these analyses to be exploratory in nature.

Results

Tables 1 and 2 provide the key features and descriptions of the independent and dependent variables for each study. The estimate of the mean effect size across the 25 studies included in the analysis using data from both standardized and unstandardized measures was 0.39 and differed significantly from zero ($p < .001$, 95% CI [0.30, 0.48]). The I^2 estimate of the percentage of between-study heterogeneity not due to chance variation in effects was 0.00%, with a τ^2 estimate of the true variance in the population of effects of 0.00. The estimate of the mean effect, the 95% confidence interval, and the heterogeneity statistics did not change when unstandardized measures were dropped from the data set. These results indicate that there was no statistically significant or meaningful heterogeneity in the study-wise effect sizes, meaning that moderator analyses were not warranted. However, for exploratory purposes, we calculated intercept-only models to determine the mean effect size at each level of the categorical moderator variables we had selected. See Table 3 for results. Additionally, a meta-regression model using pretest reading ability as a covariate predicting effect sizes across the 12 studies that included pretest data from standardized measures indicated that pretest *z* scores did not predict effect size ($\beta = -.03$, $SE = .11$, $df = 3$, $p = .78$). However, RVE estimates from models with fewer than 4 degrees of freedom are unreliable (Tipton, 2015), making the results for this analysis inconclusive.

Lastly, we explored year of publication as a covariate predicting effect size across all 25 studies. Results from the meta-regression model indicated that year of publication did not predict effect size ($\beta = -.01$, $SE = .01$, $df = 12$, $p = .20$).

Publication Bias

Because unpublished studies were not included in this meta-analysis, publication bias threatens the validity of our results, and its impact must be explored. Using the trim-and-fill method (Duval & Tweedie, 2000), we evaluated the potential impact of publication bias based on a random effects model. The trim-and-fill method removes effect sizes that cause asymmetry in a funnel plot of the effect sizes included in the meta-analysis, calculates a mean effect, and then imputes the effect sizes needed to make the plot symmetrical. The results indicate how many studies may be missing from the meta-analysis due to publication bias and produces an adjusted effect size based on including the missing studies. Results of the trim-and-fill analysis for the present meta-analysis found that publication bias might have inflated the mean effect size estimate. Seven studies with effect sizes that were smaller than the mean effect of 0.39 likely were missing from the data set. When effect sizes from these missing studies were included, the adjusted mean effect size was 0.28 (95% CI [0.20, 0.37]).

Discussion

The purpose of the current study was to conduct a meta-analysis of the research on intensive early reading interventions (i.e., provided for 100 or more sessions) for students in Grades K through 3. Specifically, we sought to update and extend Wanzek and Vaughn's (2007) synthesis of intensive interventions to allow examination of a larger corpus of studies from 1995 to 2005. We examined intervention and student characteristics that may be related to student outcomes to provide educators and researchers with the current state of evidence on the implementation of intensive interventions. We also examined whether there were systematic changes in effects of these studies since the previous synthesis.

Effects of Intensive Early Reading Interventions

The weighted mean effect size estimate (ES = 0.39), with a mean effect size adjusted for publication bias (ES = 0.28), both significantly different from zero, suggests intensive early reading interventions result in positive outcomes for early struggling readers in kindergarten through third grades. Students in these grades receiving intensive interventions like the ones in this meta-analysis may make improvements in reading by as much as four-tenths of a standard deviation. The effects of these intensive early reading interventions are similar to those reported for less intensive early reading interventions of less than 100 sessions (Wanzek et al., 2016). These findings suggest a variety of early reading interventions can improve student reading outcomes. The study of less intensive interventions did yield significant variance in effects among the studies, but it was not explained by intervention type, instructional group size, grade level, implementer, or total hours of intervention. The corpus of intensive early reading intervention did not yield significant variance in effects, further suggesting these interventions can assist students in improving their reading outcomes. We discuss the common elements of these studies that may have led to the lack of significant variance in the effects.

Over 90% of the effect sizes contributing to the weighted mean effect size estimate in the current study were calculated from results on standardized measures; this enhances our

confidence that the findings were associated with meaningful gains in reading as standardized measures are more representative of generalized reading skill than proximal, specialized measures that are often aligned to the intervention (Lipsey et al., 2012). These effect sizes are also larger than typical effects for elementary academic interventions when outcomes are measured on standardized measures, which range from 0.08 for broad subject matter measures to 0.25 for more narrow standardized measures on specific components of reading (Lipsey et al., 2012). The previous intensive reading intervention synthesis did not report a weighted mean estimate but reported study-wise mean effects ranging from −0.05 to 0.84 (Wanzek & Vaughn, 2007).

We hypothesized that publication year would significantly predict effect size, suggesting the effects of interventions conducted more recently may yield less impact, a finding reported in previous research for reading interventions with older students (Scammacca et al., 2015). However, our moderator analysis indicated publication year did not significantly predict effect size, suggesting there is no difference in the impact of intensive interventions across the 20-year time span. One reason for this might be that nearly all of the measures included in this corpus were standardized. Scammacca et al. (2015) also reported no moderation of publication year when only standardized measures were included in the moderation model. It appears that across time, intensive interventions like those included in this meta-analysis result in positive, generalized reading improvement for early struggling readers.

Characteristics of Intensive Interventions Related to Student Outcomes

We attempted to identify features of the interventions (e.g., group size, hours of intervention) included in this meta-analysis that were associated with increased effectiveness; however, we did not find significant heterogeneity in effect sizes across studies. The predominant use of standardized measures may be one reason for the lack of heterogeneity across study effects. We did calculate weighted mean effects for the intervention characteristics. We recognize that the results of these analyses may differ in a future data set that has a larger sample of studies with more heterogeneity of variance in the corpus of effect sizes.

The success of multitiered systems of support is predicated on the provision of increasingly intensive interventions in response to students' needs. One suggestion for intensifying interventions when students respond limitedly to a less intensive intervention is by reducing group size (Vaughn et al., 2012). In the previous synthesis, Wanzek and Vaughn (2007) reported effects for one-on-one instruction (*M* ES range, 0.17–0.84) were generally larger than effects for group instruction (*M* ES range, –0.05 to 0.39), but no statistical evaluation of reducing group size could be conducted. Consistent with those findings, this meta-analysis provides some limited evidence that students may respond more favorably to interventions provided in one-on-one settings (ES = 0.59) versus groups (i.e., two to eight students; ES = 0.33). One reason for this finding may be that one-on-one instruction allows for heightened response (i.e., more frequent feedback, data-based instructional changes) to a student's instructional needs that may be more difficult to achieve with more than one student in a group. Also, most of the studies utilizing one-on-one instruction were with first graders, and thus, it is not possible to know whether this age group is more responsive to one-on-one instruction or the effects are simply due to higher growth of first graders in these

interventions. Unfortunately, the research base on intensive interventions does not yet have enough studies utilizing small groups (e.g., two to four students) to examine these smaller groups versus larger instructional groups, a problem noted in the earlier synthesis as well.

Another way to increase intervention intensity is by increasing the duration of the intervention (Vaughn et al., 2012). While all of the included studies had 100 or more sessions, the length of time for each session varied. In the previous review, the authors were unable to examine the effect of intervention hours due to limited information provided in the included articles. Results from the studies reported in the current meta-analysis suggest more studies reported information needed to determine the hours of intervention. The finding that interventions provided for more than 63 hours as well as 25 to 63 hours produced significant moderate effects suggests that providing an extended number of sessions may be more important than the number of hours of intervention.

Overall, the results of this meta-analysis indicate that intensive interventions result in positive gains in reading performance for struggling readers in Grades K through 3. There is limited variability in the effects, indicating that intervention commonalities may be driving the positive effects more than differences such as group size or duration. Intervention commonalities across studies included the following: (a) a high level of standardization in which all students received the same instruction using a set of well-prescribed lessons and materials for modeling and guiding students in learning new reading practices; (b) instructional content addressing phonological awareness (e.g., syllable segmentation, phoneme identification and manipulation), phonics and word recognition (e.g., letter-name and letter-sound correspondence, blending and segmenting the sounds in words, reading decodable words and high-frequency words), and fluency (e.g., initial reading, rereading, and shared reading of decodable texts); and (c) school staff or community members implemented the interventions. Thus, generally standardized, explicit instruction including reading foundational skills provided for more than 100 sessions has a positive effect for students with reading difficulties in Grades K through 3. These interventions can be feasibly implemented by school personnel.

Implications, Limitations, and Future Research

The effective implementation of multitiered systems of support provides increasingly intensive intervention in response to student needs. This meta-analysis provides further evidence for school-based teams to consider as they make decisions about interventions for students with significant reading problems in grades kindergarten through third grade. This review provides a strong rationale for providing early intervention with the most effective instructional practices.

This meta-analysis synthesizes causal evidence for the effects of intensive reading intervention on the reading out-comes of students in Grades K through 3 with reading difficulties. However, future research is needed to enhance our understanding of intensive interventions. One-on-one instruction may improve students' response to intervention in primary grades but may not necessarily be more effective than small group instruction, which was not implemented in enough studies for examination. Additional research examining small group instruction is needed to better ascertain the effect of intensity of

intervention based on group size. Researchers might consider examining the long-term implications of one-on-one and small group instruction. Do the long-term benefits of one-on-one and small group instruction differ? Given the cost and resource demands associated with one-on-one instruction, it may be more economically and practically feasible to implement small group intervention, particularly if the long-term benefits are similar.

Additionally, researchers might consider investigating the effects of intensive interventions specifically for second-and third-grade students. The previous synthesis reported 13 studies investigating interventions that began in kindergarten and first grades and 5 studies investigating interventions that began in second and third grades. All of the studies conducted since 2005 ($k = 13$) addressed struggling readers in Grades K and 1. As such, there is limited evidence available for the impact of intensive interventions on second- and third-grade students. Given that students in Grades 2 and 3 may present more challenging reading difficulties, it is important to explore the most effective intervention for these learners. Furthermore, the majority of the more recent studies provided instruction in phonological awareness, phonics and word recognition, and fluency. Only five studies included comprehension instruction (Denton et al., 2010; Harn, Linan-Thompson, & Roberts, 2008; Little et al., 2012; Simmons et al., 2007; Vaughn et al., 2006). Struggling readers in Grades 2 and 3 may benefit from multicomponent interventions that include explicit comprehension and vocabulary components in addition to word recognition and fluency.

The current meta-analysis also highlighted the continued lack of research on individualized interventions. Researchers might examine the impact of nonstandardized approaches compared to standardized interventions. Standardized interventions are defined as a set of well-prescribed lessons and materials. The level of standardization may be high, such that all students in the group receive the same instruction, or low, meaning that interventionists still provide the same lessons to all students but may make minor adjustments based on student levels (Wanzek & Vaughn, 2007). In contrast to standardized interventions, an individualized or problem-solving approach provides instruction based on student needs. This involves identifying the problem and potential causes, creating a plan to address the problem, implementing the plan, and then evaluating the plan. This instructional process may also be referred to as data-based individualization. As with the previous corpus (i.e., 1995–2005), we identified no kindergarten through third-grade studies conducted since 2005 that examined an individualized approach to intensive intervention. It is important to identify the relative impact of standardized versus problem-solving approaches.

Due to a limited number of studies ($k = 12$) reporting standardized pretest scores in reading, we were unable to fully examine pretest reading ability as a predictor of intervention effectiveness. Our findings, though not considered reliable due to the small number of studies, indicated no differences in effects of the intensive interventions based on students' incoming levels of achievement. We recommend that researchers report scores on standardized pretest measures so that the role of initial reading level in treatment effects can be more fully explored across studies. Understanding intervention effects related to specific reading levels at pretest may allow educators to better match intensive intervention to students with identified reading levels prior to receiving Tier 3 type interventions.

Acknowledgments

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A150269 and the Eunice Kennedy Shriver National Institute of Child Health and Human Development Grant P50 HD052117. The opinions expressed are those of the authors and do not represent views of the Institutes, the U.S. Department of Education, or the National Institutes of Health.

References

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(5), 300–316.

Borenstein M, Hedges LV, Higgins JPT, & Rothstein HR (2013). Comprehensive meta analysis [software] Version 3.3.070 Englewood, NJ: Biostat.

*Coyne MD, Little M, Rawlinson D, Simmons D, Kwok O, Kim M, & Civetelli C (2013). Replicating the impact of a supplemental beginning reading intervention: The role of instructional context. Journal of Research on Educational Effectiveness, 6, 1–23. doi:10.1080/19345747.2012.706694

*Coyne MD, Simmons DC, Hagan-Burke S, Simmons LE, Kwok OM, Kim M, . . . Ware S (2013). Adjusting beginning reading intervention based on student performance: An experimental evaluation. Exceptional Children, 80, 25–44.

Denton CA, Fletcher JM, Anthony JL, & Francis DJ (2006). An evaluation of intensive intervention for students with persistent reading difficulties. Journal of Learning Disabilities, 39, 447–466. [PubMed: 17004676]

*Denton CA, Nimon K, Mathes PG, Swanson EA, Kethley C, Kurz TB, & Shih M (2010). Effectiveness of a supplemental early reading intervention scaled up in multiple schools. Exceptional Children, 76, 394–416.

Duval S, & Tweedie R (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. Biometrics, 56, 455–463. [PubMed: 10877304]

*Fien H, Smith JLM, Smolkowski K, Baker SK, Nelson NJ, & Chaparro E (2015). An examination of the efficacy of a multitiered intervention on early reading outcomes for first grade students at risk for reading difficulties. Journal of Learning Disabilities, 48, 602–621. [PubMed: 24532827]

Fisher Z, & Tipton E (2015). Robumeta: An R-package for robust variance estimation in meta-analysis Retrieved from arXiv preprint arXiv:1503.02220

*Foorman BR, Francis DJ, Winikates D, Mehta P, Schatschneider C, & Fletcher JM (1997). Early interventions for children with reading disabilities. Scientific Studies of Reading, 1, 255–276.

Francis DJ, Shaywitz SE, Stuebing KK, Shaywitz BA, & Fletcher JM (1996). Developmental lag versus deficit models of reading disability: A longitudinal individual growth curves analysis. Journal of Educational Psychology, 88, 3–17.

Fuchs D, Compton DL, Fuchs LS, Bryant J, & Davis NG (2008). Making “secondary intervention” work in a three-tier responsiveness-to-intervention model: Findings from the first-grade longitudinal reading study of the National Research Center on Learning Disabilities. Reading and Writing, 21, 413–436. doi:10.1007/s11145-007-9083-9

Fuchs D, Fuchs LS, & Vaughn S (2014). What is intensive instruction and why is it important? Teaching Exceptional Children, 46(4), 13–18.

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

*Gunn B, Biglan A, Smolkowski K, & Ary D (2000). The efficacy of supplemental instruction in decoding skills for Hispanic and Non-Hispanic students in early elementary school. The Journal of Special Education, 34, 90–103.

Hagan-Burke S, Kwok OM, Zou Y, Johnson C, Simmons D, & Coyne MD (2011). An examination of problem behaviors and reading outcomes in kindergarten students. The Journal of Special Education, 45, 131–148. doi:10.1177/0022466909359425

*Harn BA, Linan-Thompson S, & Roberts G (2008). Intensifying instruction: Does additional instructional time make a difference for the most at-risk first graders? Journal of Learning Disabilities, 41, 115–125. doi:10.1177/0022219407313586 [PubMed: 18354932]

*Hatcher PJ, Hulme C, Miles JN, Carroll JM, Hatcher J, Gibbs S, . . . Snowling MJ (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. [PubMed: 16898996]

Hedges LV, Tipton E, & Johnson MC (2010). Robust variance estimation in meta-regression with dependent effect size estimates. Research Synthesis Methods, 1, 39–65. doi:10.1002/jrsm.5 [PubMed: 26056092]

Hernandez DJ (2011). Double jeopardy: How third-grade reading skills and poverty influence high school graduation Retrieved from http://files.eric.ed.gov/fulltext/ED518818.pdf

Institute of Education Sciences. (2008). What Works Clearinghouse study review standards Retrieved from http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_version1_standards.pdf

*Jenkins JR, Peyton JA, Sanders EA, & Vadasy PF (2004). Effects of reading decodable texts in supplemental first-grade tutoring. Scientific Studies of Reading, 8, 53–85.

Juel C (1988). Learning to read and write: A longitudinal study of 54 children from first to fourth grades. Journal of Educational Psychology, 80, 437–447.

Lam EA, & McMaster KL (2014). Predictors of responsiveness to early literacy intervention: A 10-year update. Learning Disability Quarterly, 37, 134–147. doi:10.1177/0731948714529772

*Little ME, Rawlinson D, Simmons DC, Kim M, Kwok OM, Hagan-Burke S, . . . Coyne MD (2012). A comparison of responsive interventions on kindergarteners' early reading achievement. Learning Disabilities Research & Practice, 27, 189–202.

Lipsey MW, Puzio K, Yun C, Hebert MA, Steinka-Fry K, Cole MW, . . . Busick MD (2012). Translating the statistical representation of the effects of education interventions into more readily interpretable forms (NCSER 2013–3000). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education.

Lou Y, Abrami PC, Spence JC, Poulsen C, Chambers B, & d'Apollonia S (1996). Within-class grouping: A meta-analysis. Review of Educational Research, 66, 423–458.

*MacDonald C, & Figueredo L (2010). Closing the gap early: Implementing a literacy intervention for at-risk kindergartners in urban schools. The Reading Teacher, 63, 404–419.

McNamara JK, Scissons M, & Gutknectch N (2011). A longitudinal study of kindergarten children at risk for reading disabilities: The poor really are getting poorer. Journal of Learning Disabilities, 44, 421–430. doi:10.1177/0022219411410040 [PubMed: 21772059]

*Mathes PG, Denton CA, Fletcher JM, Anthony JL, Francis DJ, & Schatschneider C (2005). The effects of theoretically different instruction and student characteristics on the skills of struggling readers. Reading Research Quarterly, 40, 148–182.

*Miller SD (2003). Partners-in-reading: Using classroom assistants to provide tutorial assistance to struggling first-grade readers. Journal of Education for Students Placed at Risk, 8, 333–349.

*Morris D, Tyner B, & Perney J (2000). Early steps: Replicating the effect of a first-grade reading intervention program. Journal of Educational Psychology, 92, 681–693.

National Center on Intensive Intervention. (2012). Intensifying interventions for struggling students through data-based individualization in academics Retrieved from https://intensiveintervention.org/video-resource/intensifying-interven-tions-struggling-students-through-data-based-individualization

Nelson RJ, Benner GJ, & Gonzalez J (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review. Learning Disabilities Research and Practice, 18, 255–267.

Partanen M, & Siegel LS (2014). Long-term outcome of the early identification and intervention of reading disabilities. Reading and Writing, 27, 665–684. doi:10.1007/s11145-013-9472-1

*Santa CM, & Hoien T (1999). An assessment of early steps: A program for early intervention of reading problems. Reading Research Quarterly, 34, 54–79.

Scammacca N, Roberts G, Vaughn S, & Stuebing K (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. Journal of Learning Disabilities, 48, 369–390. doi: 10.1177/0022219413504995 [PubMed: 24092916]

*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.

*Simmons DC, Coyne MD, Hagan-Burke S, Kwok OM, Simmons L, Johnson C, . . . Crevecoeur YC (2011). Effects of supplemental reading interventions in authentic contexts: A comparison of kindergarteners' response. Exceptional Children, 77, 207–228.

Simmons DC, Coyne MD, Kwok O, McDonagh S, Harn BA, & Kame'enui EJ (2008). Indexing response to inter-vention: A longitudinal study of reading risk from kindergarten through third grade. Journal of Learning Disabilities, 41, 158–173. doi:10.1177/0022219407313587 [PubMed: 18354935]

*Simmons DC, Kame'enui EJ, Harn B, Coyne MD, Stoolmiller M, Santoro LE, . . . Kaufman NK (2007). Attributes of effective and efficient kindergarten reading intervention: An examination of instructional time and design specificity. Journal of Learning Disabilities, 40, 331–347. [PubMed: 17713132]

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

Swanson HL, Hoskyn M, & Lee C (1999). Interventions for students with learning disabilities: A meta-analysis of treatment outcomes New York, NY: Guilford Press.

Tipton E (2015). Small sample adjustments for robust variance estimation with meta-regression. Psychological Methods, 20, 375–393. doi:10.1037/met000001 [PubMed: 24773356]

Torgesen JK, Alexander AW, Wagner RK, Rashotte CA, Voeller KS, & Conway T (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. Journal of Learning Disabilities, 34, 33–58, 78. [PubMed: 15497271]

*Torgesen JK, Wagner RK, Rashotte CA, 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, 40–56. [PubMed: 20052566]

*Torgesen JK, Wagner RK, Rashotte CA, Rose E, Lindamood P, & Conway T (1999). Preventing reading failure in young children with phonological processing dis-abilities: Group and individual responses to instruction. Journal of Educational Psychology, 91, 579–593.

*Vadasy PF, Jenkins JR, Antil LR, Wayne SK, & O'Connor RE (1997). Community-based early reading intervention for at-risk first graders. Learning Disabilities Research & Practice, 12, 29–39.

*Vadasy PF, Sanders EA, & Peyton JA (2005). Relative effectiveness of reading practice or word-level instruction in supplemental tutoring: How text matters. Journal of Learning Disabilities, 38, 364–380. [PubMed: 16122070]

*Vadasy PF, Sanders EA, Peyton JA, & Jenkins JR (2002). Timing and intensity of tutoring: A closer look at the conditions for effective early literacy tutoring. Learning Disabilities Research & Practice, 17, 227–241.

Vaughn S, Denton CA, & Fletcher JM (2010). Why intensive interventions are necessary for students with severe reading difficulties. Psychology in the Schools, 47, 432–444. doi:10.1002/pits [PubMed: 21072127]

Vaughn S, Elbaum BE, Wanzek J, Scammacca N, & Walker MA (2014). Code sheet and guide for education-related intervention study syntheses Austin, TX: The Meadows Center for Preventing Educational Risk.

Vaughn S, & Linan-Thompson S (2003). What is special about special education for students with learning disabilities. The Journal of Special Education, 37, 140–147.

Vaughn S, Linan-Thompson S, & Hickman P (2003). Response to instruction as a means of identifying students with reading/ learning disabilities. Exceptional Children, 69, 391–409.

Vaughn S, Linan-Thompson S, Kouzekanani K, Bryant DP, Dickson S, & Blozis SA (2003). Grouping for reading instruction for students with reading difficulties. Remedial and Special Education, 24, 301–315.

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

Vaughn S, & Wanzek J (2014). Intensive interventions in reading for students with reading disabilities: Meaningful impacts. Learning Disabilities Research and Practice, 29, 46–53. [PubMed: 24910504]

Vaughn S, Wanzek J, Murray CS, & Roberts G (2012). Intensive interventions for students struggling in reading and mathematics: A practice guide Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Vaughn S, Zumeta R, Wanzek J, Cook B, & Klingner JK (2014). Intensive interventions for students with learning disabilities in the RTI era: Position statement of the Division for Learning Disabilities Council for Exceptional Children. Learning Disabilities Research and Practice, 29, 90–92.

Vellutino FR, Scanlon DM, Small S, & Fanuele DP (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. Journal of Learning Disabilities, 29, 157–169.

*Wang C, & Algozzine B (2008). Effects of targeted interven-tion on early literacy skills of at-risk students. Journal of Research in Childhood Education, 22, 425–439. doi: 10.1080/02568540809594637

Wanzek J, & Vaughn S (2007). Research-based implications from extensive early reading interventions. School Psychology Review, 36, 541–561.

Wanzek J, Vaughn S, Scammacca N, Gatlin B, Walker MA, & Capin P (2016). Meta-analyses of the effects of Tier 2 type reading interventions in grades K–3. Educational Psychology Review, 28, 551–576. [PubMed: 27594774]

Wanzek J, Vaughn S, Scammacca NK, Metz K, Murrary CS, Roberts G, & Danielson L (2013). Extensive reading interventions for students with reading difficulties after grade 3. Review of Educational Research, 83, 163–195. doi:10.3102/0034654313477212

Wanzek J, Wexler J, Vaughn S, & Ciullo S (2010). Reading interventions for struggling readers in the upper elementary grades: A synthesis of 20 years of research. Reading and Writing, 8, 889–912. doi:10.1007/s11145-009-9179-5

Table 1.

Features of Intervention.

Study	*n*	Grade	Frequency	Duration (in weeks)	Group Size	Implementer
Coyne, Little, et al. (2013)	162 AR	K	Daily, 30 minutes	25	3–5	General education teacher
Coyne, Simmons, et al. (2013)	103 AR	K	Daily; 30 minutes	25	3–5	MI
Denton et al. (2010)	422 AR/LD/Disab	1	Daily; 40 minutes	25	3–4	General education teacher
Fein et al. (2015)	267 AR	1	Daily; 30 minutes	26	3–5	MI
Foorman et al. (1997)	114 LD	2 & 3	Daily; 60 minutes	36	8	Special education teacher
Gunn, Biglan, Smolkowski,and Ary (2000)	256 SR	K to 3	Daily; 25–30 Minutes	60–69	1–3	Researcher
Hagan-Burke et al. (2011)	206 AR	K	Daily; 30 minutes	21	3–5	MI
Harn, Linan-Thompson, and Roberts (2008)	54 AR	1	Daily; 30 or 60 Minutes	24–25	3–5	MI-R
Hatcher et al. (2006)	77 SR	K	Daily; 20 minutes	10 or 20	1–3	Paraprofessional
Jenkins, Peyton, Sanders, and Vadasy (2004)	99 SR	1	4× week; 30 Minutes	25	1	Paraprofessional
Little et al. (2012)	90 LD/Disab[a]	K	Daily; 30 minutes	16–27	5	Reading interventionist
Mathes et al. (2005)	389 AR	1	Daily; 40 minutes	35	3	Researcher
Miller (2003)	65 SR	1	4× week; 40 Minutes	36	1	Paraprofessional
Morris, Tyner, and Perney (2000)	86 SR	1	Daily; 30 minutes	~32	1	MI
Santa and Hoien (1999)	49 SR	1	Daily; 30 minutes	35	1	MI
Schneider, Roth, and Ennemoser (2000)	253 AR	K	Daily; 10–20 minutes	10–20	5–8	General education teacher
Simmons et al. (2007)	112 AR	K	Daily; 30 minutes	21.6	5	MI
Simmons et al. (2011)	206 AR	K	Daily; 30 minutes	20	3–5	MI
Torgesen et al. (1999)	180 SR	K and 1	4× week; 20 minutes	130	5–6	MI
Torgesen, Wagner, Rashotte, Herron, and Lindamood (2010)	112 AR	1	4× week; 50 minutes	24–26	3	Researcher
Vadasy, Sanders, and Peyton (2005)	57 Disab	1	4× week; 30 minutes	32	6–8	Paraprofessional
Vadasy, Jenkins, Antil, Wayne, and O'Connor (1997)	35 AR	1	4× week; 30 minutes	27	1	Community volunteers
Vadasy, Sanders, Peyton, and Jenkins (2002)	65 AR	1 & 2	4× week; 30 minutes	35–70	1	Paraprofessional
Vaughn et al. (2006)	48 AR	1	Daily; 50 minutes	18–28	3–6	Researcher
Wang and Algozzine (2008)	139 AR	1	NR; 10–15 minutes	NR	NR	Paraprofessional

Note.AR = at risk; Disab = disabilities, type not specified; LD = learning disabilities; MI = multiple school-level implementers; NR = not reported; SR = struggling readers; MI-R = multiple school-level implementers and researchers.

[a]Sample included several disability types.

Table 2.

Description of Interventions.

Study	Measures
Coyne, Little, et al. (2013) T (ERI): Instruction in PA and writing and spelling with previously taught phonics skills C (Harcourt Trophies): Typical school intervention using phonics, guided reading, and PA	WRMT/NU (letter name and sound checklists, WA, WID), PPVT, DIBELS (PSF and NWF), CTOPP (Sound Matching, Blending Words)
Coyne, Simmons, et al. (2013) T1 (ERI): Instruction in PA, writing, and spelling with previously taught phonics skills T2 (ERI with adjustments): Implemented ERI but used in-program assessments to measure mastery then systematically target specific skills, regrouped students, and made instructional changes	WRMT/NU (Letter Identification, WA, Letter Sound Checklist, PC), PPVT-III, CTOPP (Sound Matching, Blending Words), DIBELS (PSF, NWF, ORF), Test of Written Spelling
Denton et al. (2010) T (responsive reading instruction): Instruction in PA, phonics, WR, fluency, and comprehension C (typical school practice): Typical reading instruction	WJ-III (LWID, WA, PC, Spelling), CTOPP (Blending Words, Segmenting Words), TOWRE (SWE, PDE), Comprehensive Monitoring of Early Reading Skills (ORF)
Fein et al. (2015) T (Tier 2 intervention): Instruction in PA, phonics/WR, and fluency C (business as usual): Standard district program	DIBELS (NWF, ORF), WRMT (Basic Skills Cluster, Total Reading), SAT 10
Foorman et al. (1997) T1 (synthetic phonics): based on Orton-Gillingham approach; multisensory instruction in PA and phonics T2 (analytic phonics): scripted instruction in phonics and fluency T3 (Edmark reading program): Instruction in WR, fluency, and spelling	Orthographic processing, phonological analysis, and word Reading
Gunn, Biglan, Smolkowski, and Ary (2000) T (reading mastery or corrective reading): Explicit instruction in PA, letter-sound correspondence, and blending; students placed based on reading level C (Control): No supplemental instruction	WJ-R ACH (WA, LWID, PC, Reading Vocabulary), DIBELS (ORF)
Hagan-Burke et al. (2011) T (explicit, systematic, code-based instruction): Instruction in PA, writing and spelling (school-designated instruction): Typical school instruction	CTOPP (Sound Matching, Blending Words), WRMT-R (Letter Name Checklist, Letter Sound Checklist, WA, WID)
Harn, Linan-Thompson, and Roberts (2008) T1 (less intensive intervention): Instruction in phonics and WR, fluency, passage reading, and comprehension; 30-minute sessions T2 (more intensive intervention): Instruction in phonics and WR, fluency, passage reading, and comprehension; 60-minute sessions	WRMT-R (WA, WID, PC), TOWRE (SWE, PDE), DIBELS (NWF, ORF)
Hatcher et al. (2006) T1 (modified sound-linkage reading intervention): Group instruction in letter ID, PA, and writing; individual instruction in phonics and fluency for 20 weeks T2 (modified sound-linkage reading intervention): Group instruction in letter identification, PA, and writing; individual instruction in phonics and fluency for 10 weeks	British Picture Vocabulary Test, Phonological Abilities Test, Sound Linkage Test of Phonological Awareness, Letter Identification, Early Word Recognition Test, British Ability Scales, Word Reading Test
Jenkins, Peyton, Sanders, and Vadasy (2004) T1 (sound partners–more decodable texts): Instruction in PA, phonics, WR, and spelling using more decodable texts T2 (sound partners–less decodable Texts): Instruction in PA, phonics, WR, and spelling using less decodable texts C (typical practice): Typical classroom instruction	PPVT-R, CTOPP (RLN, NWR), Yopp-Singer Segmentation test, Modified Rosner's deletion test;,WRMT-R (WA), Diagnostic Test of Basic Decoding Skills, WRAT-R (Reading, Spelling), WRMT-R (WID), TOWRE, Phonetically controlled passage (fluency, accuracy),[a] Non- phonetically controlled passage (fluency, accuracy)[a]

Study	Measures
Little et al. (2012) T (ERI): Instruction in PA and writing and spelling with previously taught phonics skills; student progress assessed every 4 weeks to determine adequate progress and make adjustments C (school-designated instruction): Typical school instruction	DIBELS (PSF, NWF, ORF), WRMT-R (Letter Name and Sounds checklist, WID, WA, Letter ID), PPVT-III, CTOPP (Sound Matching, Blending Words), Test of Written Spelling-4
Mathes et al. (2005) T1 (proactive reading): Explicit instruction in phonics/WR (in isolation), fluency, and comprehension T2 (reactive reading): Instruction in phonics/WR (not in isolation), fluency, and comprehension C (enhanced instruction): Teachers were given progress monitoring data and trained to use data to inform their typical instruction	WJ-III (WA, WID, PC, Spelling, Reading Fluency), CRAB-R (Fluency, Comprehension)
Miller (2003) T1 (partner in reading): Instruction in phonics/WR and comprehension T2 (Reading Recovery): Instruction in phonics/WR and comprehension C (school-designated instruction): Typical classroom instruction	Metropolitan Achievement Test (Word Recognition, Comprehension, Vocabulary, Language)
Morris, Tyner, and Perney (2000) T (Early Steps): Instruction in phonics/WR, fluency, and comprehension C (school-designated instruction): Typical classroom instruction	WRMT (WA, PC), Word Recognition,[a] Spelling,[a] Passage Reading[a]
Santa and Hoien (1999) T (Early Steps): Instruction in phonics/WR, fluency, and comprehension C (school-designated instruction): Typical classroom instruction	Spelling,[a] Word Recognition,[a] Early Steps Passage Reading
Schneider, Roth, and Ennemoser (2000) T1 (PA training): Instruction in PA and phonics/WR T2 (PA and letter sound training): Instruction in PA and phonics/WR T3 (letter sound training): Instruction in letter sound correspondence	Lundberg and Wimmer Metalinguistic Battery (Initial Sound Analysis, Identification of End Sounds, New phoneme Analysis, Word Length Analysis, Supply of Initial Consonant, Vowel Substitution), Wurzburger Leise Leseprobe test, Weingartener Basic Vocabulary Spelling Test: Diagnostic Spelling Test for Second Graders
Simmons et al. (2007) T1 (highly specified design): Systematic instruction in PA, phonics, writing, and spelling T2 (highly specified design + intervention): Systematic instruction in PA, phonics, writing, and spelling with instruction in comprehension T3 (moderately specified design): Instruction in PA, phonics, writing, and spelling	PPVT, DIBELS (Letter Name Fluency, PSF, NWF), Tangel & Blachman Spelling Test, Yopp Singer Test (Phoneme Segmentation), WRMT-R (WA, WID), Letter Dictation Fluency
Simmons et al. (2011) T (ERI): Instruction in PA and writing and spelling with previously taught phonics skills C (school-designated instruction): Typical Kindergarten instruction	WRMT-R/NU (WID, Letter Name Checklist, Letter Sound Checklist, PC), DIBELS (NWF, PSF), Test of Written Spelling-4, TOWRE (PDE, SWE), CTOPP (Sound Matching, Blending Words)
Torgesen et al. (1999) T1 (PA plus synthetic phonics): Explicit instruction in phonics and PA T2 (embedded phonics): Less intensive phonics instruction T3 (regular classroom support): Individual tutoring based on regular classroom instruction C (no treatment): No additional tutoring	WRMT (WA, WID, PC), WRAT (Spelling), GORT- III (Comprehension), Developmental Spelling WJPB Calculation
Torgesen, Wagner, Rashotte, Herron, and Lindamood (2010) T1 (Read, Write, and Type): Direct instruction in PA and phonics/WR with computer practice T2 (Lindamood phoneme sequencing program): Direct instruction in PA and phonics/WR with computer practice C (no treatment): No additional instruction	WRMT (WID, WA, PC), TOWRE (SWE, PDE), GORT- III (Reading Accuracy, Text Reading Fluency), CTOPP (BW, Segmenting Words, Rapid Naming Digits and Letters), Developmental Spelling Analysis, Stanford Binet Intelligence Scale 4th edition (Vocabulary), WRAT (Spelling)
Vadasy, Sanders, and Peyton (2005) T1 (reading practice): Intervention time split between instruction in phonics/WR and fluency T2 (word study): Entire intervention time spent on phonics/WR C (school-designated instruction) Typical classroom instruction	WRMT-R/NU (WA, WID, PC), WRAT-R (Reading, Spelling), TOWRE (PDE, SWE), Passage Reading Fluency, Passage Reading Accuracy

Study	Measures
Vadasy, Jenkins, Antil, Wayne, and O'Connor (1997) T (intervention): Instruction in PA, phonics/WR, and fluency C (school designated instruction): Typical classroom instruction	Analytical Reading Inventory, Dolch Word Recognition Test, WRAT-R (Reading, Spelling), WJ-R (WA, WID), Yopp-Singer Segmentation Task, Writing Sample, Bryant Pseudo-Word Test
Vadasy, Sanders, Peyton, and Jenkins (2002) T1 (sound partners and thinking partners): Instruction in PA, phonics/WR, fluency, and comprehension T2 (thinking partners): Instruction in comprehension C (school-designated instruction): Typical classroom instruction	WRAT-R (Reading, Spelling), WRMT (WID, WA), TOWRE (PDE, SWE), Read Naturally Comprehension,[a] Informal Reading Inventory
Vaughn et al. (2006) T (reading and language development intervention): Instruction in PA, phonics/WR, fluency, oral language/vocabulary, and comprehension C (school-designated instruction): Typical classroom instruction	WLPB-R (Picture Vocabulary, Verbal Analogies, Memory for Sentences, Oral Language Composite, WA, Dictation, PC), DIBELS (Letter Name Identification, ORF, Letter Sound Identification), CTOPP (RLN, NWR)
Wang and Algozzine (2008) T (supplemental targeted intervention): Direct instruction in PA, phonics/WR, and fluency C (school-designated instruction): Typical classroom instruction	WRMT-R (WID, WA, PC), DIBELS (PSF, NWF)

Note. T = treatment; C = comparison; ERI = early reading intervention; CRAB-R = Comprehensive Assessment of Reading Battery Revised for First-Grade; CTOPP = Comprehensive Test of Phonological Processing; DIBELS = Dynamic Indicators of Basic Early Literacy Skills; GORT-III = Gray Oral Reading Test III; LWID = Letter Word Identification; NWF = Nonsense Word Fluency; NWR = Non-word Repetition task; ORF = Oral Reading Fluency; PA = phonological awareness; PC = Passage Comprehension; PSF = Phoneme Segmentation Fluency; SWE = Sight Word Efficiency; PDE = Phonemic Decoding Efficiency; PPVT = Peabody Picture Vocabulary Test; RLN = Rapid Letter Naming, SAT 10 = Stanford Achievement Test–10th edition; TOWRE = Test of Word Reading Efficiency; WA = Word Attack; WID = Word Identification; WJ-III = Woodcock Johnson III; WJPB = Woodcock-Johnson Psycho-Educational Battery; WJ-R = Woodcock Johnson-Revised; WJ-R ACH = Woodcock Johnson Test of Achievement; WLPB-R = Woodcock Language Proficiency Battery–Revised; WMRT/NU = Woodcock Reading Mastery Tests–Revised/Normative Update; WR = word recognition; WRMT = Woodcock Reading Mastery Tests; WRMT-R = Woodcock Reading Mastery Tests–Revised.

[a]Unstandardized measure.

Table 3.

Effect Size by Categorical Variables.

Variable	Coefficient	*SE*	95% CI	*p*	*df*	I^2	τ^2	*n*	k
Hours of treatment									
63 or less	0.33	0.07	0.17, 0.50	.001	11	25.74	.07	167	12
More than 63	0.45	0.05	0.34, 0.56	<.001	9	0	0	145	12
Group size									
Small group	0.33	0.05	0.23, 0.43	<.001	15	2.92	.01	220	17
One-on-one	0.59	0.06	0.45, 0.73	<.001	10	0	0	108	8

Note. *k* = number of studies; *n* = number of effect sizes.