



Contents lists available at ScienceDirect

Journal of School Psychology

journal homepage: www.elsevier.com/locate/jschpsyc

Meta-analysis of targeted small-group reading interventions

Matthew S. Hall^a, Matthew K. Burns^{b,*}^a University of Minnesota, United States^b University of Missouri, United States

ARTICLE INFO

Action Editor: Steve Kilgus

ABSTRACT

Small-group reading interventions are commonly used in schools but the components that make them effective are still debated or unknown. The current study meta-analyzed 26 small-group reading intervention studies that resulted in 27 effect sizes. Findings suggested a moderate overall effect for small-group reading interventions (weighted $g = 0.54$). Interventions were more effective if they were targeted to a specific skill ($g = 0.65$), then as part of a comprehensive intervention program that addressed multiple skills ($g = 0.35$). There was a small correlation between intervention effects and group size ($r = 0.21$) and duration ($r = 0.11$). Small-group interventions led to a larger median effect size ($g = 0.64$) for elementary-aged students than for those in middle or high school ($g = 0.20$), but the two confidence intervals overlapped. Implications for research and practice are discussed.

1. Introduction

For students to effectively learn, instruction must match their diverse levels and needs (Al Otaiba & Fuchs, 2006; Kamps & Greenwood, 2005). Small-group intervention is defined as supplemental instruction delivered simultaneously to three or more students with homogenous skills to support their reading needs (Gersten et al., 2009). Small-group interventions provide the opportunity for students with reading difficulties to receive literacy instruction that more closely matches their needs. To date, small-group intervention is the component of a reading response-to-intervention model with the strongest research base (Gersten et al., 2009).

Previous research indicates that small-group reading interventions should (a) focus on the five areas of reading instruction (phonemic awareness, phonics, fluency, vocabulary, and comprehension) as outlined by the National Reading Panel (2000), (b) be implemented three to five times per week for approximately 20 to 40 min each session, and (c) build skills gradually while providing opportunity for frequent interventionist-student interaction (Gersten et al., 2009). Although small-group reading interventions are a common method used with struggling readers in U.S. schools (Foorman & Torgesen, 2001), important questions remain regarding the practice. We will describe the research, as well as evidence gaps or areas of inconsistency, in the literature below.

1.1. Effects of small-group reading interventions

Research has consistently demonstrated the positive effects of small-group reading interventions with students at risk for reading failure in early elementary (Kamps et al., 2008; Nielsen & Friesen, 2012), upper elementary (Faggella-Luby & Wardwell, 2011), middle school (Faggella-Luby & Wardwell, 2011; Vaughn et al., 2011), and high school (Bemboom & McMaster, 2013). Previous meta-analytic research found large effects for interventions with struggling readers in grades 4 through 12 ($g = 0.95$; Scammacca et al., 2007), but an update of the meta-analysis found smaller effects ($g = 0.49$) and hypothesized that the change was due to use of

* Corresponding author at: 109 Hill Hall, University of Missouri, Columbia, MO 65211, United States.

E-mail address: burnsmk@missouri.edu (M.K. Burns).

<https://doi.org/10.1016/j.jsp.2017.11.002>

Received 22 June 2016; Received in revised form 19 October 2017; Accepted 8 November 2017

Available online 15 November 2017

0022-4405/ © 2017 Society for the Study of School Psychology. Published by Elsevier Ltd. All rights reserved.

more standardized measures, differences in participant characteristics, more rigorous research designs, and improvements in the business-as-usual control group (Scammacca, Roberts, Vaughn, & Stuebing, 2015).

1.1.1. Standardized measures

Researchers have long recognized that standardized norm-referenced measures of reading are less sensitive to growth and individual differences in reading than briefer measures such as curriculum-based measurement (Marston, Fuchs, & Deno, 1986). Although standardized measures correlate well with state accountability tests (Shapiro & Gebhardt, 2012), curriculum-based measurement results in much more sensitive data (Miura Wayman, Wallace, Wiley, Tichá, & Espin, 2007). Scammacca et al. (2015) compared the effects from standardized measures of reading to the data from a previous meta-analysis and found smaller average effect size ($g = 0.21$) as compared to the previous research ($g = 0.42$; Scammacca et al., 2007). Thus, the result of using standardized measures to assess the effect of reading interventions appears to be an area in need of additional research.

1.1.2. Student characteristics

Beyond outcome measures, the age of the student receiving the small-group intervention could also influence the results. Previous research has consistently found that age of the student affected the results of reading intervention, but the difference was most pronounced for students in preschool and kindergarten (Suggate, 2016), and the effects of reading interventions for students in kindergarten maintained through the third grade (Simmons et al., 2008). Moreover, a synthesis of previous research found that reading interventions were most effective for students in kindergarten and first grade, but the review only included students through third grades (Wanzek & Vaughn, 2007). Vaughn et al. (2010) found small effects ($d = 0.16$) for a small-group reading intervention with students in middle school. However, meta-analytic research found larger mean effects for students in 4th through 12th grades (Scammacca et al., 2007, 2015). Such equivocality in findings to date suggests the grade of students receiving small-group reading intervention is also an area in need of additional research.

1.1.2.1. Research designs. When reviewing the research of a specific topic such as reading interventions, study methodology and quality of the research design should be taken into account to reduce potential bias in the findings (Wortman, 1994). Higher quality designs include studies that use randomized assignment and have low attrition or equal groups after attrition (What Works Clearinghouse, 2008). Systematic reviews of the reading intervention literature found that randomized experiments are a common method of small-group reading intervention research (Slavin, Lake, Davis, & Madden, 2011). Previous research is somewhat conflicting in that while some studies have indicated study quality did not influence reading outcomes (Ehri et al., 2001), other investigations have suggested higher quality research designs yielded larger effects (Elbaum, Vaughn, Tejero Hughes, & Watson Moody, 2000; Piasta & Wagner, 2010). In addition, research has demonstrated that the quality of the control or comparison group needs to be considered, as this group represents the counterfactual against which intervention effects are evaluated and estimated (Lemons, Fuchs, Gilbert, & Fuchs, 2014).

1.2. Characteristics of the intervention

There are other areas of potential research for small-group interventions because many components of what makes small-group reading interventions effective are still debated or unknown (Begeny, Krouse, Ross, & Mitchell, 2009; Chambers et al., 2011; Slavin et al., 2011). Below we will discuss the intervention group size, who administers the intervention, and intervention duration as characteristics of the intervention that could influence the effects.

1.2.1. Interventionist

Delivering interventions to multiple students simultaneously can be more efficient and require less time than when administering them to each student separately, but who is administering the intervention is also important to consider when determining intervention efficiency. In an extensive review of the literature on effective reading interventions, Slavin et al. (2011) found that interventions delivered by paraprofessionals and volunteers were effective, but the interventions were more effective when delivered by certified teachers. Moreover, volunteers can effectively implement reading interventions, but the effectiveness may depend on how well the volunteers were trained (Elbaum et al., 2000).

1.2.2. Intervention dosage

Another important consideration for intervention efficiency is the length of time the intervention is implemented. Researchers are more frequently considering intervention dose, or the number of teaching episodes per intervention session, which is cumulatively computed by multiplying the number of teaching episodes per session, by the number of sessions per week, by the number of weeks (Warren, Fey, & Yoder, 2007). Most intervention researchers do not report teaching episodes, but report total intervention duration in minutes or hours as an indicator of dose. It may seem intuitive that the longer an intervention is delivered the more effective it will be, but meta-analytic research found that intervention length was not associated with reading outcomes (Elbaum et al., 2000; Swanson, 1999). There could be a nonlinear relationship between intervention duration and effects because moderate-length interventions had the largest effects, and interventions that were short or relatively long in duration were equally effective (Ehri et al., 2001).

1.2.3. Group size

The size of the group remains a question in need of research. Previous research found that small-group interventions were at least as effective as one-on-one interventions (Ehri et al., 2001; Elbaum et al., 2000; Piasta & Wagner, 2010). Vaughn et al. (2003) directly compared the impact of different intervention group sizes by administering the same reading intervention to students individually, in groups of 3, and in groups of 10. Interventions delivered in groups of 3 students were more effective than groups of 10, and equally effective to the individually delivered interventions, but there was little guidance beyond comparing 3 to 10 students. Moreover, meta-analytic research found a small relationship ($r = -0.09$) between effects and size of the group (Suggate, 2016).

1.3. Targeted interventions

As described above, research regarding the effects of small-group interventions has found mixed results due to various factors, one of which could be the target of the intervention. A standardized small-group intervention that addressed word recognition, vocabulary, fluency, and comprehension led to only small effects for struggling readers ($d = 0.16$, Vaughn et al., 2010). There were several hypotheses regarding why the intervention led to only small effects. For example, the Vaughn et al. (2010) study was conducted at a middle school, and most reading intervention research is conducted with elementary-aged students, but meta-analytic research with adolescent struggling readers found average effect sizes that ranged from $g = 0.49$ (Scammacca et al., 2015) to $g = 0.95$ (Scammacca et al., 2007). Perhaps the interventions in the Vaughn et al. (2010) study were not targeted enough to address the specific reading deficits of students in the group.

Burns and colleagues (Burns & Coolong-Chaffin, 2006; Burns & Gibbons, 2012; VanDerHeyden & Burns, 2010) defined targeted interventions as those interventions that focused most directly on one area of the National Reading Panel (NRP; National Institute of Child Health and Human Development, 2000), according to student needs. Research supported targeted interventions based on the five areas of NRP because phonological decoding predicted word reading, and the rate and accuracy of word reading predicted comprehension among struggling readers (Berninger, Abbott, Vermeulen, & Fulton, 2006). Moreover, Suggate (2016) found moderate effects for interventions that focused on phonemic awareness ($d = 0.47$), phonics ($d = 0.50$), and comprehension ($d = 0.58$), and Ehri et al. (2001) found an average large effect ($d = 1.38$) from 35 studies that examined small-group reading interventions for phonemic awareness.

The targeted approach to implementing small-group intervention was tested by Burns et al. (2016) with 631 s- and third-grade students as part of the Path to Reading Excellence in School Sites (PRESS Research Team, 2014) project. Interventions delivered in small groups were targeted according to phonemic awareness, decoding, reading fluency, or vocabulary and comprehension, which was compared to a control group that received a comprehensive small-group intervention that focused on comprehension and reading fluency, with decoding and phonemic awareness embedded within the lessons. The targeted intervention led to significantly more growth on measures of reading fluency and comprehension than the control group, with moderate to large effects ($\eta^2 = 0.12$ for second grade and $\eta^2 = 0.16$ for third grade). Although the finding was consistent with previous intervention research that found interventions were more effective if they correctly targeted the student's area of challenge (Burns, VanDerHeyden, & Zaslofsky, 2014), the study was quasi-experimental with little control over the intervention that students received in the control group. Thus, additional research is needed to determine the potential differential effects for targeted small-group interventions.

1.4. Purpose

Although small-group reading interventions are commonly used, many components of what makes small-group reading interventions effective are still debated. With the large number of students with reading difficulties, the limited resources available in schools, and the increasing popularity of multi-tiered systems of support that utilize small-group instruction, there is a need to gain a better understanding of how to most effectively and efficiently deliver small-group reading interventions. Moreover, meta-analytic methods are effective means to synthesize literature in which potentially contradicting findings occur and to make specific recommendations for practice (Kavale & Forness, 2000). Therefore, the goal of the current study was to apply meta-analytic techniques to examine the effects of targeted small-group reading interventions and to identify components of small-group instruction that were most effective.

This meta-analysis was guided by the following research questions:

1. How effective are small-group reading interventions at increasing reading skills?
2. How effective are small-group reading interventions when they target specific skills as compared to a more comprehensive approach that addresses multiple skills.
3. How do intervention characteristics such as interventionist, dose, and group size impact the effect of small-group reading interventions?
4. What effect do participant-level variables (e.g., student grade) have on the effectiveness of small-group reading interventions?
5. How do study characteristics such as research design, outcome measure used, and control group impact the results of small-group reading intervention studies?

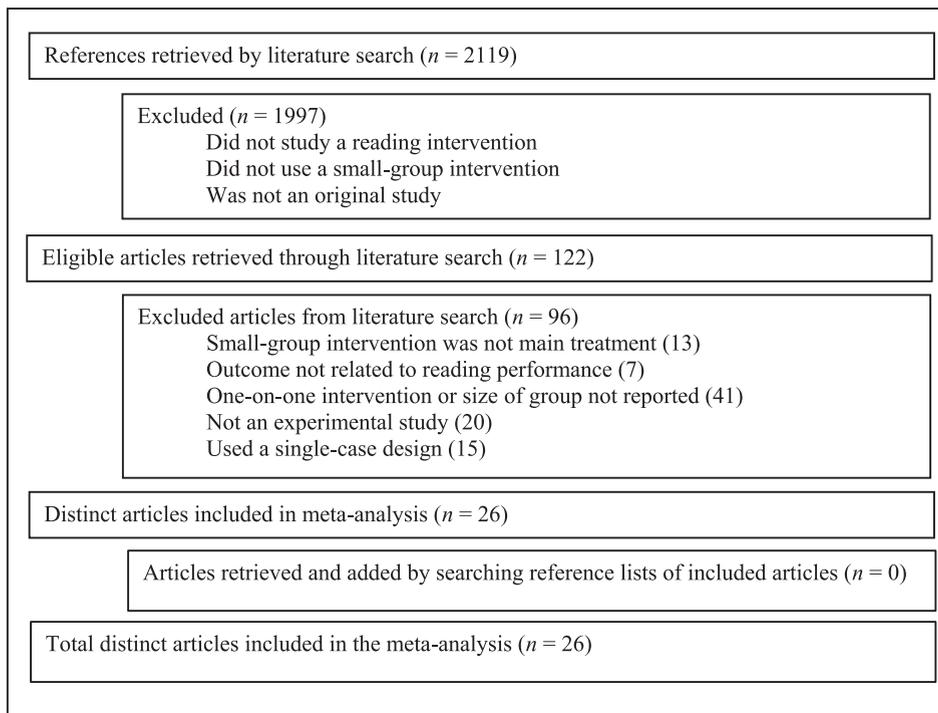


Fig. 1. Flow chart of the study selection process.

2. Method

2.1. Data collection

The PsycINFO, ERIC, and Education Full Text electronic databases were searched for articles using the terms “reading or literacy,” “small group” “tier II or tier 2,” “phonemic awareness,” “decoding,” “fluency,” “vocabulary,” “comprehension,” and “intervention.” As shown in Fig. 1, the electronic search resulted in 2119 citations, the titles of which were reviewed for relevance. A total of 1903 were eliminated without further review for lack of relevance. A total of 216 citations were further reviewed by examining their abstracts, and another 94 citations were eliminated because they were not directly about small-group reading interventions, leaving 122 total articles. The following inclusionary criteria were used to compare the remaining articles in order to find data usable for the current meta-analysis:

1. The study implemented a small-group reading intervention.
2. The outcome variables of the study were directly related to reading or literacy performance.
3. The group size of the intervention included between 3 and 15 students. Three was established as the minimum to avoid research with dyads, and because it was used as the minimum group size in previous research (Vaughn et al., 2003). A group size of 15 was set as the maximum to avoid research with entire classrooms.
4. The intervention was delivered by an adult rather than peer to peer (e.g., peer learning groups or cooperative learning groups).
5. The participants were students in grades K through 12.
6. The participants were native English speakers.
7. The study was published in a peer-reviewed journal.
8. The study used a quantitative group comparison research design, either experimental or quasi-experimental, allowing for estimates of effect size as compared with a control condition.
9. The study presented sufficient information to calculate effect sizes.
10. The study was written in English.

Thirteen of the 122 articles were excluded because a small-group reading intervention was not used as a main treatment (e.g., the reading intervention was only one part of a multi-component intervention), 7 were eliminated because the outcome variables were not directly related to reading performance, 41 were excluded because they used group sizes that were less than three students or did not give the group size, and 20 were removed because they were a review or did not use a quantitative research design. Studies using a single-case design method were not included in the current study due to the difficulty of interpreting effect sizes and combining the effects with group design studies (Baron & Derenne, 2000), which resulted in 15 articles being excluded.

After identifying articles appropriate for the meta-analysis from the electronic searches, an ancestral search of the reference lists from included articles was conducted to identify potential additional articles, but no additional studies were found. In total, 26 articles met the inclusion criteria and were included in the meta-analysis. It should be noted that one of the articles reported two independent studies within the same article (i.e., Study 1 and Study 2; Coyne, McCoach, & Kapp, 2007), which were coded separately and increased the number of studies to 27.

The interrater agreement for whether the studies met the inclusion criteria was assessed by the first author and a school psychology graduate student by randomly selecting 25% ($n = 26$) of the articles initially screened. The total agreements were divided by the total articles that were randomly selected and initially resulted in 81% agreement. All of the disagreements involved the size of the groups in the study, where the second rater included studies that used groups of two students. The two raters met to discuss the disagreements and reached 100% agreement after clarifying the intervention group size variable.

2.2. Coding

The 27 articles were coded according to variables relevant to the research question. The coding process is described below.

2.2.1. Targeted interventions

Three studies used either a phonemic awareness or phonics intervention (e.g., teaching the sounds that corresponded to individual phonemes; O'Shaughnessy & Swanson, 2000), one study used a reading fluency intervention (repeated reading and reading widely; Kuhn, 2005), five studies used a vocabulary intervention (explicit instruction with contextual information and definitions using multiple expositors to the word; Coyne et al., 2007), and four addressed reading comprehension (e.g., teaching students how to summarize, question, clarify, and predict as they read; Sporer, Brunstein, & Kieschke, 2009). Thus, 13 studies targeted the intervention to one of the areas identified by the NRP and were coded as targeted interventions. An additional 14 studies (including both presented by Coyne et al., 2007) used an intervention that included a combination of aforementioned intervention type categories, which were coded as comprehensive interventions.

2.2.2. Intervention variables

A variety of intervention variables were coded to determine how intervention characteristics affected the results of small-group reading interventions. The study examined who implemented the intervention (i.e., interventionist), the intervention dose, and intervention group size.

2.2.2.1. Interventionist. Three groups of treatment agents, or those who implemented the intervention, were used in the included studies. Researchers or graduate student research assistants administered the small-group intervention for 14 of the studies; personnel from within the school (e.g., teachers or paraprofessionals) implemented the intervention for 10 of the studies; and 4 effects were from studies that used volunteers or other trained interventionists from outside of the school to provide the small-group interventions.

2.2.2.2. Intervention dose. Treatment dose was coded as total time in hours the intervention was delivered. Minutes were converted to a decimal point by dividing the minutes by 60 (e.g., 90 min equaled 1.5 h). There was considerable variability in intervention time with a range of 1.5 h to 1440 h over the course of several weeks. The mean number of hours was 123.59 ($SD = 287.00$) and the median was 30.00. Three studies did not provide information regarding the amount of time dedicated to the intervention.

2.2.2.3. Intervention group size. Intervention group size was coded as a continuous count of the total number of students in each instructional group. The size of the intervention groups varied greatly between the studies and many studies used a range of group sizes within the same study. The median group size was used for studies reporting a range of group sizes. The average group size was 4.98 ($SD = 2.41$) students and the median was 4.50.

2.2.3. Participant variable: grade

Studies meeting the criteria for inclusion were coded according to the grade of the participants in their study. Nineteen studies included students who were in kindergarten through 5th grade and were classified as elementary. Seven studies included students who were in 6th through 12th grade and were classified as secondary. One study included students in both groups.

2.2.4. Study variables

Finally, variables associated with study characteristics were coded. The variables of interest were research design, outcome measure used, and type of control group.

2.2.5. Research design

In order to account for the quality of the research design, studies were coded into three categories using methods from the [What Works Clearinghouse \(2008\)](#) Procedures and Standards. The research standards take into account if the study was an experimental or quasi-experimental design, the amount of attrition in the study, and the equality of the groups for studies of high attrition. Fifteen effects were from studies that were categorized as meeting the standards (i.e., study included a randomized design with low attrition and equality across groups), seven were from studies that met the standards with reservations (i.e., quasi-experimental designs or was

a randomized design with either high attrition or unequal groups), and five were from studies that did not meet the standards (i.e., did not use randomization with high attrition and unequal groups).

2.2.5.1. Outcome measure. The outcome measure used was coded to examine the effect that use of standardized measure had on intervention effectiveness. Studies that used standardized norm-referenced measures were classified as standardized, and those that used researcher-developed measures or curriculum-based measurement were coded as informal measures. A total of 8 studies used standardized measures (e.g., *Test of Word Reading Efficiency*; Torgesen, Rashotte, & Wagner, 1999), 9 used informal measures (e.g., curriculum-based measures and researcher-developed comprehension questions), and 10 used a combination of the two. Although not used in the analyses, the type of measure used was further coded according to reading skill measured and reported in Table 1. One study only measured phonemic awareness (measures of blending and segmenting sounds; Kerins, Trotter, & Schoenbrodt, 2010). One study only measured phonics (measures of letter-sound knowledge, word recognition, and spelling; Schuele et al., 2008). One study measured reading fluency (word reading efficiency; Kuhn, 2005). Five studies examined the effect on vocabulary (e.g., expressive measure of story word definitions, receptive measure of story word definitions, and receptive measure of understanding story words in context; Coyne et al., 2007). Four studies used comprehension as the dependent variable (cloze, strategy use, and group-administered standardized measure of reading comprehension; Faggella-Luby & Wardwell, 2011). Finally, 15 studies used a combination of measures to examine the effect on reading.

2.2.5.2. Control group. The intervention effect sizes were obtained in each study by comparing the effects to a control group. Therefore, the type of control group (i.e., business as usual or active) used in each study was coded to account for any impact this may have on the results. Active control groups provide some type of similar intervention to the one being studied, but is different in some important way (Boot, Simons, Stothart, & Stutts, 2013). Therefore, studies were coded as using an active control group if the control group received a different reading intervention or additional supplemental instruction. Studies were coded as using a business-as-usual control group if the control group did not receive any additional reading intervention beyond regular classroom instruction. Eighteen effects were from studies that used an active comparison group and nine were from studies that used a business-as-usual comparison group.

2.2.5.3. Fidelity. Although not included in the analysis, treatment implementation fidelity outcomes for each study were recorded. Eighteen studies reported treatment fidelity data. The mean fidelity reported ranged from 87% to 100% with an overall mean of 94% across the fifteen studies. One study (Torgesen et al., 2010) videotaped treatment sessions for fidelity and stated that treatment implementation fidelity was very high, but they did not report any data. Eight studies did not mention collecting treatment fidelity data.

2.2.6. Reliability of the coding

A second person also coded the study variables for approximately 40% ($n = 10$) of the studies. The number of variables that were coded the same way by both coders was totaled and divided by the total number of variables across the studies. The two coders agreed 100% of the time.

2.3. Effect size calculation and analysis

The study assumed random-effects and calculated effect sizes for each study with Hedges g using the formula outlined by Hedges (1981), and recommended by the What Works Clearinghouse Standards (2008), which adjusts for the size of the samples. Hedges g is a standardized mean difference statistic, calculating the difference between the post-test treatment and control group means, divided by the pooled standard deviation. Although the Hedges g index has a slight upward bias when estimating population effects, each outcome measure is adjusted to correct for this bias (Hedges, 1981). With this adjustment, Hedges g has sound statistical properties in small sample sizes.

The effect sizes for each reading measure were calculated and the mean was calculated for studies that included more than one reading measure so that each study contributed only one estimate of effect to the analyses. One study (Coyne et al., 2007) contributed two effect sizes to the meta-analysis because it reported separate studies with different samples and interventions. Effect sizes were weighted according to the inverse of the variance and median weighted effects were reported for each variable. Median effects were reported because there were too few effects for some of the variables for the data to meet the assumptions necessary for parametric statistics (i.e., mean). The 95% confidence intervals for the median effect sizes were compared to determine if the intervals overlapped. Those that did not overlap were judged to be reliably different to a $p < 0.05$ criterion. Cohen (1988) provided criteria to interpret estimates of effect for g , r , and r^2 . The criteria for g were 0.2 for a small effect, 0.5 for a medium effect, and 0.8 for a large effect. The corresponding recommended values for a small, medium, and large effect for r were 0.10, 0.30, and 0.50 respectively, and 0.02, 0.13, and 0.26 respectively for r^2 .

A failsafe N was computed (Orwin, 1983) to address potential publication biases, as only published studies were used in the meta-analysis. Failsafe N provides information on the stability of a meta-analysis by identifying how many studies with a zero effect would have to be found to change a medium or large effect to a small effect. The criterion of 0.20 was used to indicate a small effect (Cohen, 1988). The criterion for a small effect is somewhat arbitrary, but it is commonly accepted in the absence of any other criterion to guide a particular line of research.

Table 1
Summary information on interventions included in the meta-analysis by study.

Study	Grade	Group size (median)	Intervention type	Measure	Effect size (g)
Bemboom and McMaster (2013)	10	8	Multiple	Fluency Comp	0.68
Case et al. (2010)	1	3.5	Multiple	PA Phonics	0.44
Coyne et al. (2007) Study 1	K	3.5	Voc	Voc	1.64
Coyne et al. (2007) Study 2	K	3.5	Voc	Voc	1.23
Faggella-Luby and Wardwell (2011)	5–6	6	Comp	Comp	0.49
Glenberg, Brown, and Levin (2007)	1–2	3	Comp	Comp	1.99
Graves, Brandon, Duesbery, McIntosh, and Pyle (2011)	6	3	Multiple	Phonics Voc Comp	– 0.21
Graves, Duesbery, Pyle, Brandon, and McIntosh (2011)	6	3	Multiple	Phonics Voc Comp	0.22
Jimenez et al. (2010)	K – 2	5	Multiple	PA Fluency Comp	0.21
Kamps et al. (2007)	1–2	5	Multiple	Phonics Fluency Comp	1.41
Kamps et al. (2008)	K – 2	4.5	Multiple	Phonics Fluency Comp	1.52
Kuhn (2005)	2	6	Fluency	Fluency	0.07
Loftus, Coyne, McCoach, Zipoli, and Pullen (2010)	K	3.5	Voc	Voc	0.44
Mathes et al. (2005)	1	3	Multiple	PA Phonics Fluency	0.26
Mathes et al. (2003)	1	4.5	Multiple	Phonics Fluency Comp	0.90
Nielsen and Friesen (2012)	K	4.5	Voc	Voc	1.12
O'Shaughnessy and Swanson (2000)	2	5	Phonics	PA Phonics Fluency Comp	1.00
Penno, Wilkinson, and Moore (2002)	K	12	Voc	Voc	0.70
Rashotte, MacPhee, and Torgesen (2001)	1–6	4	Multiple	PA Phonics Fluency Comp	0.67
Schuele et al. (2008)	K	6	Phonics	Phonics	0.26
Sporer et al. (2009)	6	5	Comp	Comp	0.58
Torgesen, Wagner, Rashotte, Herron, and Lindamood (2010)	1	3	Multiple	PA Phonics Fluency Comp	0.51
Vaughn et al. (2010)	7–8	10.5	Multiple	Phonics Fluency Comp	0.04
Vaughn et al. (2010)	6	12.5	Multiple	PA Phonics Fluency Comp	0.17
Vaughn et al. (2011)	7–8	4.5	Multiple	PA Phonics Fluency Comp	0.26
Westerveld and Gillon (2008)	3	3.5	Comp	Comp	0.65

Note. Comp = Reading Comprehension, Fluency = Reading Fluency, PA = Phonemic Awareness, Voc = Vocabulary.

3. Results

The first research question addressed the overall effects of small-group reading interventions. A total of 27 effect sizes were calculated from 26 studies investigating the effect of small-group reading interventions on reading skills. A summary of the study

Table 2
Median effect sizes according to student and intervention variables.

Variable	<i>k</i>	Weighted <i>g</i>	95% CI	Fail safe <i>N</i>
Intervention agent				
Researcher or graduate student	14	0.51	0.18–0.84	22
Within school personnel	10	0.59	0.26–0.91	20
Outside school interventionists	3	0.52	0.31–0.73	6
Student grade				
Elementary	19	0.64	0.38–0.90	42
Secondary	7	0.20	–0.01–0.41	NA
Combination	1	0.49	NA	1
Study quality				
Meets standards	15	0.59	0.33–0.85	29
Meets with reservations	7	0.26	–0.15–0.67	2
Does not meet standards	5	0.23	–0.28–0.70	< 1
Reading outcome measure				
Standardized	8	0.39	–0.02–0.80	8
Informal	9	0.58	0.23–0.83	17
Combination	10	0.60	0.25–0.95	20
Control group type				
Business as usual	18	0.44	0.28–0.60	22
Active	9	1.32	0.89–1.75	50
Total	27	0.54	0.32–0.76	46

characteristics is presented in Table 1. The effect sizes ranged from $g = -0.21$ to 1.99 with a median weighted effect size of $g = 0.54$ (95% CI = 0.32 to 0.76).

3.1. Targeted intervention

The second research question addressed the effects of targeted small group interventions in comparison to more comprehensive intervention packages. Effect sizes were examined according to if the intervention was targeted ($k = 13$) or more comprehensive (i.e., containing components across intervention type categories, $k = 14$). Targeted interventions were found to have a larger weighted effect size ($g = 0.65$, 95% CI = 0.33 to 0.97) than comprehensive interventions ($g = 0.35$, 95% CI = 0.11 to 0.59), but the two confidence intervals overlapped. Failsafe *N* analyses found that there would have to be 29 studies with 0 effect to reduce the median effect size to fall below the criterion for a small effect (i.e., $g = 0.20$). The Failsafe *N* for comprehensive interventions suggested that 11 studies with 0 effect would have to be included to change the effect size to a small magnitude of 0.20 or less.

3.2. Intervention variables

Several different intervention variables were examined to determine which small-group reading intervention components contributed to the effect of the intervention. The results of the analyses are described below.

3.2.1. Interventionist

First, studies were grouped by the treatment agent to determine effect sizes according to the person who implemented the intervention. As shown in Table 2, interventions implemented by a researcher or graduate research assistant yielded a weighted effect size of $g = 0.51$ (95% CI = 0.18 to 0.84). Studies that used staff within the school to implement the intervention had a weighted effect size of $g = 0.59$ (95% CI = 0.26 to 0.91). Finally, interventions implemented by a trained interventionist from outside the school were found to have an effect size of $g = 0.52$ (95% CI = 0.31 to 0.73). Once again, there was considerable overlap between the three confidence intervals and the magnitudes of the weighted effect sizes were quite similar.

3.2.2. Intervention dose/duration

The dosage of the intervention was examined according to the total amount of hours the intervention was implemented. Three studies did not report the intervention amount and were not included in this analysis. The studies resulted in a mean of 123.59 ($SD = 287.00$) hours of intervention, but the data were not normally distributed (skewness = 4.37, kurtosis = 20.33). Thus, the data were normalized by identifying outliers with Tukey's interquartile range and removing the outliers from the analyses, which resulted in 2 sets of data being removed from the analysis. The resulting correlation between intervention duration and effect size was small $r = -0.11$, $p = 0.62$.

3.2.3. Intervention group size

The last intervention component that was examined was the size of the intervention groups. The mean group size was 4.98 students ($SD = 2.41$), but again the data were not normally distributed (skewness = 2.20, kurtosis = 4.94). Thus, the data were normalized by identifying outliers with Tukey's interquartile range and removing the outliers from the analyses, which resulted in 3

Table 3
Regression analyses for intervention variables and effect sizes.

	Model 1				Model 2				Model 3			
	B	S.E.	β	t	B	S.E.	β	t	B	S.E.	β	t
Constant	0.96	0.34		2.86*	1.93	0.72		2.69*	2.21	0.79		-2.80*
Targeted intervention	0.28	0.22	0.29	1.32	0.44	0.23	0.45	1.89	0.63	0.32	0.64	1.97
Group size					-0.17	0.12	-0.36	-1.51	-0.20	0.12	-0.41	-1.68
Intervention duration									0.01	0.01	0.26	0.88
	$R^2 = 0.08, F \text{ Change} = 1.74$				$R^2 = 0.19, F \text{ Change} = 2.29$				$R^2 = 0.22, F \text{ Change} = 0.78$			

* $p < 0.05$.

sets of data being removed. The resulting correlation between group size and intervention effect was small $r = -0.21, p = 0.34$.

The effects of targeted interventions, group size, and intervention dosage (number of intervention hours) were analyzed with a regression using the three variables as the predictors and intervention effect as the dependent variable. The targeted variable was coded as a 0 if it addressed multiple skills and a 1 if it was targeted to one area of reading. Data that were identified as outliers on the intervention dosage or group size data were excluded. The type of intervention was entered first, then the order of the variables was determined by the magnitude of the correlation with effect size, with group size being entered second and intervention duration third. As shown in Table 3, none of the variables were significant predictors, but type of intervention (targeted or comprehensive) accounted for a small to medium amount of variance ($r^2 = 0.08$) (Cohen, 1988). The size of the intervention group added an additional 11% variance, which is also a medium effect, but adding in intervention duration only added 3% of the variance for a total of 22%.

3.3. Participant variables: student grade

The impact of small-group reading interventions was examined by the grade of the students receiving the intervention. As shown in Table 2, the effect sizes for students in elementary school (i.e., students in grades K through 5th grades) was moderate ($g = 0.64, 95\% \text{ CI} = 0.38\text{--}0.90$), and the effect size for students in secondary grades (8th through 12th grades) was small ($g = 0.20, 95\% \text{ CI} = -0.01\text{--}0.41$). One study used students from both grade groups, and had a moderate effect ($g = 0.49$). Although the median effect size was larger for elementary than secondary students, there was some overlap between the two confidence intervals. However, the confidence interval for secondary students include zero.

3.4. Study variables

Finally, variables associated with the research were evaluated. The variables examined were quality of the research design, outcome measures used, and type of control group.

3.4.1. Research design

Studies were grouped according to the quality of the design using the categories presented in the What Works Clearinghouse (2008) Procedures and Standards. Studies that met standards yielded a weighted effect size of $g = 0.59$ ($95\% \text{ CI} = 0.33$ to 0.85). Those that met standards with reservations had a weighted effect size of $g = 0.26$ ($95\% \text{ CI} = -0.15$ to 0.67). Studies that did not meet the What Works Clearinghouse standards yielded an effect size of $g = 0.23$ ($95\% \text{ CI} = -0.28$ to 0.70). Although there was overlap between the confidence intervals, the confidence intervals for studies that met standards with reservations and those that did not meet standards both included zero.

3.4.2. Outcome measure used

Next, type of outcome measure used was examined. Standardized reading measures resulted in a median weighted effect size of $g = 0.39$ ($95\% \text{ CI} = -0.02$ to 0.80), informal measures resulted in a median weighted effect size of $g = 0.58$ ($95\% \text{ CI} = 0.23$ to 0.83), and the studies that used a combination of the two resulted in a median weighted effect size of $g = 0.60$ ($95\% \text{ CI} = 0.25$ to 0.95). There was considerable overlap between the three confidence intervals.

3.4.3. Control group

Finally, the type of control group that was used was also examined. Studies that used a business-as-usual control group were found to have a weighted effect size of $g = 0.44$ ($95\% \text{ CI} = 0.28$ to 0.60), whereas studies that used an active control group were found to have a weighted effect size of $g = 1.32$ ($95\% \text{ CI} = 0.89$ to 1.75). The two confidence intervals did not overlap, which suggested that the effect size for studies that used an active control group was reliably larger than those that used a business-as-usual condition.

4. Discussion

The current study investigated the effects of small-group reading interventions with 27 effects from 26 studies. While previous

research found that small-group reading interventions were effective (e.g., Ehri et al., 2001; Piasta & Wagner, 2010; Vaughn et al., 2003), the current study examined the effects of targeted small-group reading interventions relative to a series of variables, such as intervention-related variables, grade as a student variable, and research design.

4.1. Effectiveness of small-group interventions

An overall weighted effect size for small-group reading interventions of $g = 0.54$ was found, suggesting that small-group reading interventions in general were moderately effective. The confidence interval for the overall median effect did not include 0, and 46 studies with a 0 effect would have to be found to change the score to fall below the criterion for a small effect of $g = 0.20$. Thus, across the examined studies, small-group reading interventions reliably resulted in a positive effect for students.

4.2. Targeted small-group interventions

Interventions that were targeted to a single reading skill area were found to be more effective than more general interventions that combined multiple reading skill areas; however, the confidence intervals did overlap. One of the core elements of a multi-tiered system of support is assessing students to identify their needs and providing them with explicit instruction in those areas (Fuchs & Fuchs, 2006; Gersten et al., 2009; Justice, 2006). The current finding was consistent with previous research regarding the importance of targeted reading interventions (Burns et al., 2016) and could provide a hypothesis about why previous comprehensive small-group interventions led to small effects (Vaughn et al., 2010). Thus, providing targeted reading interventions to students in small-groups through a multi-tiered system of support or other models appears to be an effective method for increasing students' reading skills. However, additional research is needed to more directly compare the two approaches and to determine how to best identify which intervention to target for individual students.

4.3. Intervention characteristics

Several intervention characteristics were examined in the current analysis, including interventionist, dosage (as measured by total intervention duration), and group size.

4.3.1. Interventionist

The intervention agent seemed to have an impact on the effectiveness of the intervention, but there was considerable overlap between the effect size confidence intervals for all three types of treatment agents. Small-group reading interventions implemented by teachers had a small to moderate median effect size and were less effective than when the intervention was administered by researchers or by trained interventionists from outside the school. This is somewhat in contrast to previous research that found teachers to be more effective at implementing interventions than volunteers (Slavin et al., 2011) and as effective at implementing interventions as researchers (Piasta & Wagner, 2010). A possible explanation for the results in the current study is the amount of training provided to the interventionists. Studies using interventionists from outside of the school provided the interventionists with training in administering the intervention (Rashotte et al., 2001; Torgesen et al., 2010; Vaughn et al., 2010), which could account for the larger effect size and is consistent with the findings from Elbaum et al. (2000). This hypothesis would support providing adequate training to those implementing a small-group reading intervention, but additional research is needed.

4.3.2. Intervention dose

There was a small correlation between the number of intervention hours and size of the effect. Several reasons for this limited relationship are possible. For instance, certain studies might have involved interventions that were stopped once a student reached a certain criterion. In contrast, intervention dosage in other studies might have continued irrespective of student response. The shortest interventions were the vocabulary storybook reading interventions, which yielded strong effects (Coyle et al., 2007; Nielsen & Friesen, 2012; Penno et al., 2002) and suggested that small-group storybook reading interventions were an effective method for increasing young students' vocabulary in a short amount of time.

4.3.3. Group size

There was also a small and negative relationship between group size and intervention effects ($r = -0.21$). Previous research found that groups of less than five students were more effective than groups with five or more students (Vaughn et al., 2003). Given the need for schools to find more efficient ways to provide interventions to struggling learners (Fielding, Kerr, & Rosier, 2007), understanding how the size of a reading intervention group impacts the effectiveness of the intervention is an important consideration. Only one study (Vaughn et al., 2003) has systematically investigated the size of reading intervention groups, which suggests that this is an important area for future research to address.

The use of targeted interventions and group size accounted for 19% of the variance, which is a large effect (Cohen, 1988). The two variables were roughly equal in magnitude of effect ($r^2 = 0.09$ and $r^2 = 0.11$) respectively. The intervention dosage, as measured by total intervention duration, accounted for a small amount of variance. Thus, it seems that targeted interventions and group size are the two most important intervention variables among the ones examined.

4.4. Grade of students

Interventions delivered to elementary school students resulted in an effect size that was over three times as large as the one for secondary students, but the confidence intervals for the two estimates of effect did overlap. This finding is consistent with the research on literacy instruction in general that emphasizes the importance of early intervention (Good, Simmons, & Smith, 1998; Wanzek & Vaughn, 2007) and supports providing additional small-group intervention for students displaying early signs of experiencing reading difficulties. The current data are more consistent with Scammacca et al. (2015) meta-analysis that found a moderate effect ($g = 0.49$) than the older one that found a large effect ($g = 0.95$).

4.5. Study characteristics

Several study characteristics were examined in the current analysis, including the quality of the research design, the measure used, and the control group.

4.5.1. Research design

Studies that utilized high quality research designs were the most common in the present analysis and also yielded the largest overall effect size. Given that higher quality research designs are less susceptible to internal and external threats to validity and increase the confidence with which causal attributions can be made (Shadish, Cook, & Campbell, 2002), this finding is encouraging and increases the confidence that can be placed in the overall effectiveness of small-group reading interventions.

4.5.2. Outcome measure

The current study found an effect for standardized measures of $g = 0.39$, which was consistent with the two previous meta-analyses ($g = 0.21$; Scammacca et al., 2015; $g = 0.42$; Scammacca et al., 2007). The informal measures resulted in a somewhat larger effect ($g = 0.58$), but the two confidence intervals overlapped. It could be hypothesized that targeted interventions would measure outcomes with more precise and informal measures than comprehensive interventions, which could partially explain the somewhat larger effects for the targeted interventions. However, the difference between the median scores for the two types of measures was not reliably different, was not likely to be the explanation for other findings, and suggested an area in need of additional research.

4.5.3. Control group

It was somewhat surprising that studies that used an active control group had a weighted effect size that was approximately three times as large as the weighted effect for comparisons to a business-as-usual control group. Given that students in active control groups receive an intervention similar to the intervention groups, this finding is difficult to explain. Perhaps taking students out of core literacy instruction and providing an inferior intervention led to less student learning than allowing students to remain in class. Alternatively, the current data support the concept that core instruction has improved in recent years and provide a counterfactual to the research hypothesis (Lemons et al., 2014). Additional research in this area is needed given the importance of core instruction to successful reading development and the counterintuitive finding in the current study.

4.6. Limitations

The findings and conclusions of this present analysis are not without their limitations. As is the case with all meta-analyses, caution must be taken when interpreting the causal relationships of the variables. The variables were not manipulated in the current study and therefore it is not clear whether they caused the differences in effect sizes or if the results were caused by spurious factors.

Many studies that used small-group interventions were not included in the analysis because their range of intervention group size included groups of two students, and therefore did not meet the inclusion criteria of groups of 3–15 students. Consequently, the relatively small number of studies and effect sizes included in the analysis is a limitation to the findings. However, the number of studies implementing small-group reading interventions that we identified was greater than other recent extensive reviews of the reading intervention literature (e.g., Slavin et al., 2011). Moreover, we only included published studies to ensure some basic level of rigor and to avoid one data set being represented more than once. However, including only published studies could have resulted in a publication bias. We attempted to address that by including a failsafe N estimate for each variable, but the influence of only including published studies cannot be ignored.

We used the framework of the NRP (2000) to guide our search and search terms, which may be conceptually preferable to searching a wider variety of terms that could be arbitrarily selected, but may have limited the potential of the search process to identify all possible studies. Future researchers could replicate this design with even more comprehensive search terms. Moreover, we excluded peer-delivered reading interventions, which may have affected the current results. Future meta-analytic researchers could include peer-delivered interventions or study them on their own.

It would be interesting to compare the effects of small-group reading interventions on specific measures of reading. We reported the measures used in Table 1, but did not report median effect sizes for each because of the small number of studies that used some of the measures. Future meta-analytic researchers could directly examine the effect that measure has on reading interventions.

Another limitation of the current study is that components other than the small-group intervention were included in three of the studies used in the analysis. Two studies implemented a Tier-1 (i.e., classroom-wide) intervention in addition to the small-group intervention that was not provided to the comparison students and one study provided a reward component to help with behavior

during the intervention. Due to the fact that only three studies provided additional components, these were not accounted for in the analysis, but may have had an impact on the results. Finally, we used weighted effect sizes, but could not adjust the means because most studies did not include a pretest.

5. Conclusion

This meta-analysis adds to the current research demonstrating that small-group reading interventions are effective (Ehri et al., 2001; Elbaum et al., 2000; Piasta & Wagner, 2010; Scammacca et al., 2007, 2015; Vaughn et al., 2003). The results of this study suggest that reading intervention provided in small-groups was effective, and there were moderate effects for using targeted interventions. Several intervention variables were related to intervention effects, but using targeted interventions and group size were the two most closely related to the outcomes. The grade of the student and study characteristics also led to differential effects, and suggested important areas of future research.

References*

- Al Otaiba, S., & Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? *Journal of Learning Disabilities, 39*, 414–431.
- Baron, A., & Derenne, A. (2000). Quantitative summaries of single-subject studies: What do group comparisons tell us about individual performances? *Behavior Analyst, 23*, 101–106.
- Begeny, J. C., Krouse, H. E., Ross, S. G., & Mitchell, R. C. (2009). Increasing elementary-aged students' reading fluency with small-group interventions: A comparison of repeated reading, listening passage preview, and listening only strategies. *Journal of Behavioral Education, 18*, 211–228.
- *Bemboom, C. M., & McMaster, K. L. (2013). A comparison of lower- and higher-resourced tier 2 reading interventions for high school sophomores. *Learning Disabilities Research and Practice, 28*, 184–195.
- Berninger, V. W., Abbott, R. D., Vermeulen, K., & Fulton, C. M. (2006). Paths to reading comprehension in at-risk second-grade readers. *Journal of Learning Disabilities, 39*, 334–351.
- Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science, 8*, 445–454.
- Burns, M., & Gibbons, K. (2012). *Implementing response-to-intervention in elementary and secondary schools: Procedures to assure scientific-based practices* (2nd ed.). New York: Routledge.
- Burns, M. K., & Coolong-Chaffin, M. (2006). Response to intervention: The role of and effect on school psychology. *School Psychology Forum: Research in Practice, 1*, 3–15.
- Burns, M. K., Maki, E. E., Karich, A. C., Hall, M., McComas, J. J., & Helman, L. (2016). Problem-analysis at tier 2: Using data to find the category of the problem. In S. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.). *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 293–308). (2nd ed). New York: Springer.
- Burns, M. K., VanDerHeyden, A. M., & Zaslofsky, A. F. (2014). Best practices in delivering intensive academic interventions with a skill-by-treatment interaction. In P. L. Harrison, & A. Thomas (Eds.). *Best practices in school psychology: Student-level services* (pp. 129–141). Bethesda, MD: National Association of School Psychologists.
- *Case, L. P., Speece, D. L., Silverman, R., Ritchey, K. D., Schatschneider, C., Cooper, D. H., ... Jacobs, D. (2010). Validation of a supplemental reading intervention for first-grade children. *Journal of Learning Disabilities, 43*, 402–417.
- Chambers, B., Slavin, R. E., Madden, N. A., Abrami, P., Logan, M. K., & Gifford, R. (2011). Small-group, computer-assisted tutoring to improve reading outcomes for struggling first and second graders. *The Elementary School Journal, 111*, 625–640.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Lawrence Erlbaum Associates.
- *Coyne, M. D., McCoach, D. B., & Kapp, S. (2007). Vocabulary intervention for kindergarten students: Comparing extended instruction to embedded instruction and incidental exposure. *Learning Disability Quarterly, 30*, 74–88.
- Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., 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.
- Elbaum, B., Vaughn, S., Tejero Hughes, M., & Watson Moody, S. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology, 92*, 605.
- *Faggella-Luby, M., & Ardwell, M. (2011). RTI in a middle school: Findings and practical implications of a tier 2 reading comprehension study. *Learning Disability Quarterly, 34*, 35–49.
- Fielding, L., Kerr, N., & Rosier, P. (2007). *Annual growth for all students, catch-up growth for those who are behind*. Kennewick: The New Foundation Press.
- Foorman, B. R., & Torgesen, J. (2001). Critical elements of classroom and small-group instruction promote reading success in all children. *Learning Disabilities Research and Practice, 16*, 203–212.
- Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*, 93–99.
- Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2008). Assisting students struggling with reading: Response to Intervention and multi-tier intervention for reading in the primary grades. *A practice guide. (NCEE 2009-4045)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- *Glenberg, A. M., Brown, M., & Levin, J. R. (2007). Enhancing comprehension in small reading groups using a manipulation strategy. *Contemporary Educational Psychology, 32*, 389–399.
- Good, R., III, Simmons, D., & Smith, S. (1998). Effective academic interventions in the United States: Evaluating and enhancing the acquisition of early reading skills. *School Psychology Review, 27*, 45–56.
- *Graves, A. W., Brandon, R., Duesbery, L., McIntosh, A., & Pyle, N. B. (2011). The effects of tier 2 literacy instruction in sixth grade: Toward the development of a response-to-intervention model in middle school. *Learning Disability Quarterly, 34*, 73–86.
- *Graves, A. W., Duesbery, L., Pyle, N. B., Brandon, R. R., & McIntosh, A. S. (2011). Two studies of tier II literacy development: Throwing sixth graders a lifeline. *The Elementary School Journal, 111*, 641–661.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*, 107–128.
- *Jimenez, J. E., Rodriguez, C., Crespo, P., Gonzalez, D., Artiles, C., & Alfonso, M. (2010). Implementation of response to intervention (RTI) model in Spain: An example of a collaboration between Canarian universities and the Department of Education of the Canary Islands. *Psicothema, 22*, 935–942.
- 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.
- *Kamps, D., Abbott, M., Greenwood, C., Arreaga-Mayer, C., Wills, H., Longstaff, J., ... Walton, C. (2007). Use of evidence-based, small-group reading instruction for English language learners in elementary grades: Secondary-tier intervention. *Learning Disability Quarterly, 30*, 153–168.
- *Kamps, D., Abbott, M., Greenwood, C., Wills, H., Veerkamp, M., & Kaufman, J. (2008). Effects of small-group reading instruction and curriculum differences for

* Indicates included in the analysis.

- students most at risk in kindergarten: Two-year results for secondary- and tertiary-level interventions. *Journal of Learning Disabilities*, 41, 101–114.
- Kamps, D., & Greenwood, C. (2005). Formulating secondary level reading interventions. *Journal of Learning Disabilities*, 38, 500–509.
- Kavale, K. A., & Forness, S. R. (2000). Policy decisions in special education: The role of meta-analysis. In R. Gersten, E. P. Schiller, & S. Vaughn (Eds.). *Contemporary special education research: Syntheses of knowledge base on critical instructional issues* (pp. 281–326). Mahwah, NJ: Lawrence Erlbaum Associates.
- *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.
- *Kuhn, M. R. (2005). A comparative study of small-group fluency instruction. *Reading Psychology*, 26, 127–146.
- Lemons, C. J., Fuchs, D., Gilbert, J. K., & Fuchs, L. S. (2014). Evidence-based practices in a changing world: Reconsidering the counterfactual in education research. *Educational Researcher*, 43, 242–252.
- *Loftus, S. M., Coyne, M. D., McCoach, D. B., Zipoli, R., & Pullen, P. C. (2010). Effects of a supplemental vocabulary intervention on the word knowledge of kindergarten students at risk for language and literacy difficulties. *Learning Disabilities Research and Practice*, 25, 124–136.
- Marston, D., Fuchs, L. S., & Deno, S. L. (1986). Measuring pupil progress: A comparison of standardized achievement tests and curriculum-related measures. *Diagnostique*, 11, 77–90.
- *Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). The effects of theoretically different instruction and student characteristics on the skills of struggling readers. *Reading Research Quarterly*, 40, 148–182.
- *Mathes, P. G., Torgesen, J. K., Clancy-Menchetti, J., Santi, K., Nicholas, K., Robinson, C., & Grek, M. (2003). A comparison of teacher-directed versus peer-assisted instruction to struggling first-grade readers. *The Elementary School Journal*, 103, 459–479.
- Miura Wayman, M., Wallace, T., Wiley, H. I., Tichá, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *Journal of Special Education*, 41, 85–120.
- National Institute of Child Health and Human Development (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769)*. Washington, DC: U.S. Government Printing Office.
- *Nielsen, D. C., & Friesen, L. D. (2012). A study of the effectiveness of a small-group intervention on the vocabulary and narrative development of at-risk kindergarten children. *Reading Psychology*, 33, 269–299.
- Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics*, 8, 157–159.
- *O'Shaughnessy, T. E., & Swanson, H. L. (2000). A comparison of two reading interventions for children with reading disabilities. *Journal of Learning Disabilities*, 33, 257–277.
- *Penno, J. F., Wilkinson, I. A. G., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology*, 94, 23–33.
- Piasta, S. B., & Wagner, R. K. (2010). Developing early literacy skills: A meta-analysis of alphabet learning and instruction. *Reading Research Quarterly*, 45, 8–38.
- PRESS Research Team (2014). *PRESS intervention manual*. Minneapolis, MN: Minnesota Center for Reading Research, University of Minnesota.
- *Rashotte, C. A., MacPhee, K., & Torgesen, J. K. (2001). The effectiveness of a group reading instruction program with poor readers in multiple grades. *Learning Disability Quarterly*, 24, 119–134.
- Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C. K., & Torgesen, J. K. (2007). *Interventions for adolescent struggling readers: A meta-analysis with implications for practice*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.
- Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *Journal of Learning Disabilities*, 48, 369–390.
- *Schuele, C. M., Justice, L. M., Cabell, S. Q., Knight, K., Kingery, B., & Lee, M. W. (2008). Field-based evaluation of two-tiered instruction for enhancing kindergarten phonological awareness. *Early Education and Development*, 19, 726–752.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Shapiro, E. S., & Gebhardt, S. N. (2012). Comparing computer-adaptive and curriculum-based measurement methods of assessment. *School Psychology Review*, 41, 295–305.
- Simmons, D. C., Coyne, M. D., Kwok, O. M., McDonagh, S., Harn, B. A., & Kame'enui, E. J. (2008). Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities*, 41, 158–173.
- Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, 6, 1–26.
- *Sporer, N., Brunstein, J. C., & Kieschke, U. (2009). Improving students' reading comprehension skills: Effects of strategy instruction and reciprocal teaching. *Learning and Instruction*, 19, 272–286.
- 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, H. L. (1999). Reading research for students with LD: A meta-analysis of intervention outcomes. *Journal of Learning Disabilities*, 32, 504–532.
- Torgesen, J. K., Rashotte, C. A., & Wagner, R. K. (1999). *TOWRE: Test of word reading efficiency*. Austin, TX: Pro-ed.
- *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, 40–56.
- VanDerHeyden, A. M., & Burns, M. K. (2010). *Essentials of response to intervention*. Hoboken, NJ: Wiley.
- *Vaughn, S., Cirino, P. T., Wanzek, J., Wexler, J., Fletcher, J. M., Denton, C. D., ... Francis, D. J. (2010). Response to intervention for middle school students with reading difficulties: Effects of a primary and secondary intervention. *School Psychology Review*, 39, 3–21.
- Vaughn, S., Linan-Thompson, S., Kouzekanani, K., Bryant, D. P., Dickson, S., & Blozis, S. A. (2003). Reading instruction grouping for students with reading difficulties. *Remedial and Special Education*, 24, 301–315.
- *Vaughn, S., Wanzek, J., Wexler, J., Barth, A., Cirino, P. T., Fletcher, J., ... Francis, D. (2010). The relative effects of group size on reading progress of older students with reading difficulties. *Reading and Writing*, 23, 931–956.
- *Vaughn, S., Wexler, J., Roberts, G., Barth, A. A., Cirino, P. T., Romain, M. A., ... Denton, C. A. (2011). Effects of individualized and standardized interventions on middle school students with reading disabilities. *Exceptional Children*, 77, 391–407.
- Wanzek, J., & Vaughn, S. (2007). Research-based implications from extensive early reading interventions. *School Psychology Review*, 36, 541–561.
- Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Developmental Disabilities Research Reviews*, 13, 70–77.
- *Westerveld, M. F., & Gillon, G. T. (2008). Oral narrative intervention for children with mixed reading disability. *Child Language Teaching and Therapy*, 24, 31–54.
- What Works Clearinghouse (2008). *WWC procedures and standards handbook*. Washington, DC: Institute for Education Science.
- Wortman, P. M. (1994). Judging research quality. In H. Cooper, & L. V. Hedges (Eds.). *The handbook of research synthesis* (pp. 97–109). New York: Russell Sage Foundation.