

# A Meta-Analysis of Interventions for Struggling Readers in Grades 4–12: 1980–2011

Journal of Learning Disabilities  
2015, Vol. 48(4) 369–390  
© Hammill Institute on Disabilities 2013  
Reprints and permissions:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/0022219413504995  
journaloflearningdisabilities.sagepub.com



Nancy K. Scammacca, PhD<sup>1</sup>, Greg Roberts, PhD<sup>1</sup>,  
Sharon Vaughn, PhD<sup>1</sup>, and Karla K. Stuebing, PhD<sup>2</sup>

## Abstract

This meta-analysis synthesizes the literature on interventions for struggling readers in Grades 4 through 12 published between 1980 and 2011. It updates Scammacca et al.'s analysis of studies published between 1980 and 2004. The combined corpus of 82 study-wise effect sizes was meta-analyzed to determine (a) the overall effectiveness of reading interventions studied over the past 30 years, (b) how the magnitude of the effect varies based on student, intervention, and research design characteristics, and (c) what differences in effectiveness exist between more recent interventions and older ones. The analysis yielded a mean effect of 0.49, considerably smaller than the 0.95 mean effect reported in 2007. The mean effect for standardized measures was 0.21, also much smaller than the 0.42 mean effect reported in 2007. The mean effects for reading comprehension measures were similarly diminished. Results indicated that the mean effects for the 1980–2004 and 2005–2011 groups of studies were different to a statistically significant degree. The decline in effect sizes over time is attributed at least in part to increased use of standardized measures, more rigorous and complex research designs, differences in participant characteristics, and improvements in the school's "business-as-usual" instruction that often serves as the comparison condition in intervention studies.

## Keywords

struggling readers, reading intervention, reading disabilities

Results from the 2011 National Assessment of Educational Progress indicate that just 34% of both fourth graders and eighth graders are reading at or above a proficient level (National Center for Education Statistics, 2011). One third of fourth graders and nearly one fourth of eighth graders scored below basic in their reading proficiency, meaning that they lack the ability to comprehend text written at their grade level. These data highlight the fact that many students in Grade 4 and above require reading intervention to improve their comprehension skills. These skills are critical for acquiring content knowledge from what they read. Without effective intervention, they will lack the literacy skills needed to join the workforce or pursue postsecondary education (Kamil et al., 2008). However, as Kamil et al. (2008) point out, educators of students in Grades 4 to 12 often struggle to know how best to help students improve their reading ability

In response to this need for research-based guidance on selecting and implementing reading interventions beyond the primary grades, Scammacca et al. (2007) conducted a meta-analysis of the empirical literature published between 1980 and 2004 on interventions for struggling readers in Grades 4 through 12. The aim of the meta-analysis was to

determine the relative effectiveness of these interventions and provide evidence-based guidance for policy and practice. Scammacca et al. found an overall mean effect size of 0.95 across all types of reading interventions and outcome measures, indicating that on average, the group receiving treatment outperformed the comparison group by nearly one standard deviation. The mean effect size when looking at outcomes on standardized, norm-referenced measures only was far smaller at 0.42, or nearly half a standard deviation of difference between treatment and comparison groups. Researchers found similar results when looking at the effects of reading interventions on measures of reading comprehension, reporting a 0.97 mean effect for all measures of reading comprehension and 0.35 mean effect for standardized measures of reading comprehension. These effects are similar to those reported by Edmonds et al.

<sup>1</sup>University of Texas at Austin, USA

<sup>2</sup>University of Houston, TX, USA

## Corresponding Author:

Nancy K. Scammacca, University of Texas at Austin, 1 University Station  
D4900, Austin, TX 78712, USA.

E-mail: nancyscammacca@austin.utexas.edu

(2009), who calculated a mean effect size of 0.89 on measures of reading comprehension from 13 reading intervention studies involving students in Grades 6 to 12 published between 1994 and 2004.

In subsequent moderator analyses, Scammacca et al. (2007) examined the extent to which these mean effect sizes varied based on the type of intervention, the grade level and learning disability (LD) status of the students who participated, and whether the intervention was implemented by a teacher or by a researcher. Results indicated that (a) 4th through 12th graders can benefit from word-level and text-level interventions, (b) students in upper elementary and middle school showed the greatest gains but older students also made measurable progress, (c) students with LD benefit from intervention when it is tailored to their needs; and (d) teachers can provide effective interventions for struggling readers. The authors concluded that future research should focus on longer interventions and measure outcomes using group-administered standardized measures.

The research base on reading interventions for students in Grades 4 to 12 has expanded considerably since the publication of Scammacca et al. (2007). As more funding became available to develop and test interventions with this population, researchers increasingly turned their attention to helping struggling readers in 4th through 12th grade. Recent intervention studies sought to extend the knowledge gained through prior research by implementing larger-scale research efforts, lengthier and more multifaceted interventions, and more rigorous research designs that utilized standardized, norm-referenced outcome measures to a greater extent than earlier studies. As a result, a sufficient number of new studies have been published to warrant a new meta-analysis. The purpose of this article is to update and expand the 2007 report with findings from studies published between 2005 and 2011 to determine whether the quality of studies reflected the guidance of the Scammacca et al. synthesis and the extent to which the conclusions drawn in the 2007 report would prevail. In addition, we intended to combine the set of studies from the 2007 report with more recent studies, allowing for sufficient statistical power to examine additional moderator variables, including the number of hours of intervention provided. Finally, we expected that comparing the more recent studies with those synthesized in the 2007 report would shed light on advancements in the rigor of research designs for literacy interventions for students in Grades 4 to 12.

### *Insights From Recent Meta-Analyses*

Since the publication of Scammacca et al. (2007), other researchers have conducted meta-analyses of selected adolescent reading interventions published in 2005 and later. Flynn, Zheng, and Swanson (2012) focused their meta-analysis on students in Grades 5 through 9 who were identified as having a reading disability. To meet their criteria,

students in the study must have scored below the 25th percentile on a standardized, norm-referenced reading measure (standard score below 90). Furthermore, the authors included results from standardized, norm-referenced measures only. The 12 studies that met the inclusion criteria yielded a mean effect size of 0.41, nearly identical to the mean effect size of 0.42 reported for standardized measures in Scammacca et al. Flynn et al. attempted to use moderator variables related to the focus and length of the intervention and student characteristics such as age and grade level to explain the statistically significant heterogeneity in their analysis. None of these variables was found to be a significant moderator. Flynn et al. concluded that the small number of studies in their meta-analysis was a significant limitation of their findings. By including a larger sample of studies, the present meta-analysis expected to overcome this limitation and identify characteristics of more and less effective interventions.

Wanzek et al. (2013) focused their meta-analysis on adolescent reading interventions that provided at least 75 sessions to students in Grades 4 through 12. Separate meta-analyses were conducted for measures of reading comprehension, reading fluency, word reading, word-reading fluency, and spelling. The number of effect sizes included in these meta-analyses ranged from five (in the meta-analysis of spelling outcomes) to 22 (in the meta-analysis of reading comprehension outcomes). Both standardized, norm-referenced measures and researcher-developed measures were included in the meta-analyses. Overall mean effect sizes ranged from 0.10 (reading comprehension outcomes) to 0.16 (reading fluency and word-reading fluency outcomes). Statistically significant heterogeneity was present only in the reading comprehension outcomes. Subsequent moderator analyses found no differences in mean effect size based on number of hours of intervention, group size, or grade level. Noting that the mean effect sizes were considerably smaller than those in earlier meta-analyses, the authors concluded that shorter interventions may be associated with larger effects than extensive interventions because the novelty of a brief intervention and/or the immediate impact of increased instructional time in reading produced an initial growth spurt that may be difficult to maintain over time. The present meta-analyses included number of hours of intervention as a moderator to determine if mean effect sizes are larger for briefer interventions than longer ones. All studies in Wanzek et al. (2013) were included in the present meta-analyses.

### *Changes Over Time in Reading Intervention Research*

Changes in federal legislation and funding priorities over the past 10 years have caused shifts in the reading intervention research landscape since the publication of the studies analyzed in Scammacca et al. (2007). The U.S. Congress

established the Institute of Education Sciences (IES) with the passage of the Education Sciences Reform Act of 2002. From its inception, one goal of IES has been to increase the rigor of education research (IES, 2005). In its 2005 Biennial Report to Congress, IES stated that its funding procedures favor rigorous research designs that emulate the types of randomized trials found in medical research. Furthermore, IES reported a 200% increase from 2001 and 2004 in the use of true experimental designs in government-funded projects. Given the time it takes to prepare a grant proposal, receive the award, carry out a study, analyze and report the results, and publish the report in a journal, it is likely that studies using rigorous experimental designs that were proposed to IES starting in 2003 would reach publication in 2005 and beyond. More rigorous research designs typically measure effects using standardized measures, which have been shown to result in lower estimates of effect size. In addition, rigorous experimental designs often implement long interventions and have large sample sizes. As a result, effect sizes from research published from 2005 onward may differ in key ways from those from earlier research.

An additional factor that introduced changes in adolescent reading intervention research was the passage of the Individuals with Disabilities Education Improvement Act (IDEIA) in 2004. IDEIA 2004 introduced a change in the procedures for identifying children in need of special education services. Schools were now permitted to identify students in need of intervention based in part on their response to instruction (RTI) instead of relying primarily finding an IQ–achievement discrepancy. RTI’s multitiered instruction approach requires that multiple levels of intensity of intervention be available to students who do not respond to classroom-level instruction. Thus, IDEIA 2004 and RTI likely affected adolescent reading intervention research published in 2005 and beyond in two ways. First, as RTI models were increasingly used to identify struggling readers, the pool of students who qualified to participate in intervention research broadened to include those who did not respond to classroom-level instruction but who either did not have a formal LD designation or who did have an LD designation but were not identified based on an IQ–achievement discrepancy. In addition, schools that were using the RTI model were implementing their own interventions for struggling readers that became the new “business-as-usual” comparison condition in intervention studies. Both of these factors could result in differences in the estimates of the effect of interventions tested from 2005 onward compared to earlier interventions.

### Research Questions

The present meta-analyses seek to replicate and update Scammacca et al. (2007) by addressing the following questions:

1. How effective are the reading interventions provided in Grades 4 to 12 that have been studied over the past 30 years, both overall and on measures of reading comprehension?
2. How does the observed magnitude of the effect of reading interventions for students in Grades 4 to 12 vary based on student, intervention, and research design characteristics?
3. Do more recently studied interventions differ from older ones in their effectiveness, both overall and on measures of reading comprehension?

## Method

### Literature Search

A computer search of ERIC and PsycINFO was conducted to locate studies published between 2005 and 2011 to add to the studies published between 1980 and 2004 that were included in Scammacca et al. (2007). The search procedure used to locate the studies for the 2007 meta-analysis was repeated. Descriptors or root forms of those descriptors (reading difficult\*, learning disab\*, LD, mild handi\*, mild disab\* reading disab\*, at-risk, high-risk, reading delay\*, learning delay\*, struggling reader, dyslex\*, read\*, comprehend\*, vocabulary, fluen\*, word, decod\*, English Language Arts) were used in various combinations to capture the greatest possible number of articles. Articles published online in 2011 in advance of their print publication were included, resulting in the inclusion of one study with a 2012 print publication date. A search of abstracts from other published research syntheses and meta-analyses was done and reference lists in seminal studies were reviewed to ensure that all relevant studies were identified. In addition, a search through all articles published between 2005 through 2011 in 11 major journals was conducted. These journals were selected because they were the journals in which previous intervention studies were published and were likely sources of high quality studies. Journals examined in this search included *Annals of Dyslexia*, *Exceptional Children*, *Journal of Educational Psychology*, *Journal of Learning Disabilities*, *Journal of Special Education*, *Learning Disabilities Research & Practice*, *Learning Disability Quarterly*, *Reading Research Quarterly*, *Remedial and Special Education*, and *Scientific Studies of Reading*.

### Inclusion Criteria

Studies found through the literature search were included in the meta-analysis if they met all of the following criteria:

1. Participants were English-speaking struggling readers. Struggling readers were defined as those with

low achievement in reading, unidentified reading difficulties, dyslexia, and/or with reading or LD. Studies also were included if disaggregated data were provided for struggling readers regardless of the characteristics of other students in the study. Only disaggregated data on struggling readers were used in the meta-analysis.

2. Participants were in Grades 4 to 12 (age 9–21). When a sample also included older or younger students and it could be determined that the sample mean age was within the targeted range, the study was accepted. Studies were included if disaggregated data were provided for students in Grades 4 to 12 even if older and/or younger students also participated in the study. Only disaggregated data on students within the targeted grade range were used in the meta-analysis.
3. The study utilized an experimental or quasi-experimental treatment-comparison or multiple-treatment comparison research design. Studies were coded as treatment-comparison designs if the comparison group received either no intervention or the school's "business-as-usual" reading intervention. Studies were coded as multiple-treatment designs if all groups received an intervention designed by researchers that they would not have received if they were not participants in the study.
4. The intervention provided any type of reading instruction, including word study, fluency, vocabulary, reading comprehension, or multiple components of reading instruction in English.
5. Data were reported for at least one dependent measure that assessed one or more reading constructs. Data from measures of other constructs, including content acquisition, reading motivation, and attitudes, were not included in the meta-analysis.
6. Sufficient data for calculating effect sizes and standard errors were provided.

The same criteria were used for determining the studies to include in the Scammacca et al. (2007) meta-analysis.

An initial search using these criteria identified 119 publications as potentially meeting all criteria. On further review, 83 publications were eliminated because they failed to meet all of the inclusion criteria. Studies most often were excluded because, on further inspection, it was determined that they did not provide a reading intervention, did not measure a reading outcome, did not use a group comparison research design, or included students in Grades 1 to 3 in the intervention along with older students and did not disaggregate results by grade level or age. The remaining 36 publications were retained for coding.

### Coding Procedures

A similar code sheet to that used by Scammacca et al. (2007) was used for coding the new studies for the present report. The code sheet included elements specified in the What Works Clearinghouse Design and Implementation Assessment Device (IES, 2008) and used in previous research (Edmonds et al., 2009; Wanzek et al., 2006). Data coded included participant characteristics, description of the methodology and intervention, indicators of study quality, properties of measures, and data needed for calculating effect sizes. Moderators of interest also were captured in the coding.

Researchers with doctorate degrees and doctoral students with experience coding studies for other meta-analyses and research syntheses completed the code sheets for Scammacca et al. (2007) and the studies added for the present meta-analysis. All coders had completed training on how to complete the code sheet and had reached a high level of reliability with others coding the same article independently. Every study included in the current meta-analysis was independently coded by two raters. When discrepancies were found between coders, they reviewed the article together and discussed the coding until consensus was reached.

### Effect Size Calculation

For all studies, the Hedges (1981) procedure for calculating unbiased effect sizes for Cohen's *d* was used (this statistic is also known as Hedges's *g*). Hedges's *g* was calculated using the posttest means and standard deviations for treatment and comparison (or multiple treatment) groups when such data were provided. In some cases, Cohen's *d* effect sizes were reported and means and standard deviations were not available. For these effects, Cohen's *d* for posttest mean differences between groups and the treatment and comparison group sample sizes were used to calculate Hedges's *g*. Sample-weighted estimates of Hedges's *g* were computed to account for potential bias in studies with small samples. All effects were computed using the Comprehensive Meta Analysis (Version 2.2.064) software (Borenstein, Hedges, Higgins, & Rothstein, 2011).

In all, 17 of the new research reports and 2 studies from the 2007 article contained more than one treatment-control or multiple-treatment group comparison. Where comparisons represented independent subgroups (consisting solely of participants whose data were not used in other comparisons in the article), effect sizes from all comparisons were entered into the meta-analysis separately. Where comparisons represented dependent subgroups (with the same participants' data represented in multiple comparisons in the article, such as when the same control group is compared to

two or more treatment groups), the procedure recommended by Borenstein, Hedges, Higgins, and Rothstein (2009) was implemented. This procedure involves computing a combined mean effect size and its variance in a manner that reflects the degree of dependence in the data. This approach differs from the procedure implemented in Scammacca et al. (2007), in which the treatment group that best represented the implementation of the intervention was included and other treatment group comparisons were dropped. As a result, one additional study that represented an independent group comparison was included and one study-wise effect size was recomputed. This resulted in some differences in mean effect sizes and confidence intervals from those provided in the original report.

Nearly all studies from Scammacca et al. (2007) and the new studies provided data on multiple outcome measures. Since these effects are inherently dependent, effect sizes from multiple measures were averaged using the procedures recommended by Borenstein et al. (2009) and the average and its standard error were included in the meta-analysis. This procedure was utilized in the 2007 report as well.

As a result of implementing the procedures described above, 82 independent, study-level effect sizes from 67 published research reports were included in the meta-analyses conducted for this article. Of these, 32 were published between 1980 and 2004 (hereafter referred to as the 1980–2004 group) and 50 were from studies published between 2005 and 2011 (hereafter referred to as the 2005–2011 group).

### Meta-analytic Procedures

A random-effects model was used to analyze effect sizes. This model allows for generalizations to be made beyond the studies included in the analysis to the population of studies from which they come. Recent methodological innovations in meta-analysis such as multilevel modeling (Hox, 2002) and structural equation modeling (Cheung, 2008) were considered as approaches to the random-effects analyses of the effect sizes. However, these models proved impossible to fit to the available data due to the number of categorical moderators of interest, many of which had more than two levels. Therefore, a traditional approach was taken to the meta-analyses. Mean effect size statistics and their standard errors were computed and heterogeneity of variance was evaluated using the  $Q$  statistic. When statistically significant variance was found, moderator variables were introduced into the random-effects models, resulting in mixed-effects models. Moderators included the following:

- Intervention type (fluency, word study, vocabulary, reading comprehension strategy, or multiple components)
- Type of implementer (teacher or researcher)

- Grade level of students (4th–5th grades, 6th–8th grades, and 9th–12th grades)
- LD status (no students with LD, some students with LD and some non-LD struggling readers, all students with LD)
- Hours of intervention provided (0–5, 6–15, 16–25, 26 or more)
- Study design (multiple treatment or treatment/comparison)

The first four moderators were selected because they were included in the Scammacca et al. meta-analysis. Hours of intervention was added to investigate the role of intensity of intervention in intervention effectiveness. Study design was selected as a moderator to determine if the comparison condition in treatment-comparison designs (typically the school's business-as-usual treatment for struggling readers in the studies located for the present meta-analysis) could be considered an alternate treatment that affected the magnitude of the effect sizes in a similar way to a researcher-designed alternative treatment.

As was done in Scammacca et al. (2007), separate meta-analyses were conducted on effect sizes from all types of measures (including measures of vocabulary, spelling, decoding, reading fluency, and reading comprehension), all standardized, norm-referenced measures of these reading skills, measures of reading comprehension, and standardized, norm-referenced measures of reading comprehension. By analyzing standardized measures separately, effects that are less closely aligned to the specific instruction provided in the intervention can be observed. Reading comprehension measures were analyzed separately because gains in comprehension generally are seen as the key goal of reading intervention.

The original analysis plan for the meta-analyses called for effect sizes from studies from the 2007 report to be combined with those from the more recent studies located for the present report. However, when the combined meta-analyses were completed, very large  $Q$  statistics were observed and the mean effects were markedly different from what had been reported in the 2007 meta-analyses. As a result, separate meta-analyses were conducted for the studies from the 2007 report (the 1980–2004 group) and the current set of studies (the 2005–2011 group) and the 95% confidence intervals for their mean effects were compared to each other to determine if they overlapped. Overlapping confidence intervals generally would suggest that the 1980–2004 and 2005–2011 sets of studies came from the same population of studies, whereas nonoverlapping confidence intervals would suggest that they come from different populations of studies (though it is possible for differences to be statistically significant even when confidence intervals overlap). Results are reported below for all studies combined, the

1980–2004 group of studies, and the 2005–2011 group of studies. See Table 1 for characteristics of all studies included in this report.

## Results

### All Types of Outcome Measures

The estimate of the mean effect size across the 82 effects from all studies was 0.49 ( $p < .001$ , 95% CI = 0.38, 0.60), indicating a moderate positive effect of intervention of nearly half a standard deviation on students' reading outcomes. The variance as measured by the  $Q$  statistic was statistically significant and very large ( $Q = 389.00$ ,  $df = 81$ ,  $p < .001$ ). The mean effect for the 1980–2004 group of studies was considerably larger at 0.96 ( $p < .001$ , 95% CI = 0.69, 1.23) than the mean effect for all studies, whereas the mean effect for the studies located for this report was somewhat smaller at 0.23 ( $p < .001$ , 95% CI = 0.15, 0.31). The confidence intervals for the two sets of studies do not overlap, indicating that the two sets do not come from the same population of studies. The result of a comparison of the mean effect sizes for the 1980–2004 and 2005–2011 sets of studies indicated that the difference was statistically significant ( $Q$ -between = 25.81,  $df = 1$ ,  $p < .001$ ). For this reason, each corpus of studies was treated separately for further analyses. The  $Q$  statistics were statistically significant for both groups of studies (for the 1980–2004 group,  $Q = 159.84$ ,  $df = 31$ ,  $p < .001$ ; for the 2005–2011 group,  $Q = 98.53$ ,  $df = 49$ ,  $p < .001$ ).

### All Standardized, Norm-Referenced Outcome Measures

The estimate of the mean effect size across the 53 effects from standardized outcomes from both sets of studies was 0.21 ( $p < .001$ , 95% CI = 0.12, 0.30), indicating a small positive effect of intervention on standardized measures of reading. The variance as measured by the  $Q$  statistic was statistically significant ( $Q = 121.15$ ,  $df = 52$ ,  $p < .001$ ). The 1980–2004 group of studies had a larger mean effect size at 0.42 ( $p = .006$ , 95% CI = 0.25, 0.59) from 12 effects from standardized outcomes. The variance as measured by the  $Q$  statistic was statistically significant ( $Q = 67.89$ ,  $df = 11$ ,  $p < .001$ ). The studies located for the present report were far more likely to include results from at least one standardized measure, with 41 of 50 (82%) contributing an effect size for this analysis compared to 12 of 32 (38%) for the 1980–2004 group. The estimate of the mean effect size for these 41 effects was 0.13 ( $p < .001$ , 95% CI = 0.07, 0.18), which is considerably smaller than the 0.42 mean effect for the 1980–2004 group of studies. The variance as measured by the  $Q$  statistic was not statistically significant for the new group of studies ( $Q = 42.03$ ,  $df = 40$ ,  $p = .38$ ). As with the

mean effects for all measures, the confidence intervals for the mean effects for standardized measures for the 1980–2004 and 2005–2011 groups of studies do not overlap, indicating that they come from different populations of studies. In addition, a comparison of the mean effect sizes for the 1980–2004 and 2005–2011 sets of studies indicated that they were different to a statistically significant degree ( $Q$ -between = 4.83,  $df = 1$ ,  $p < .028$ ).

### All Reading Comprehension Outcome Measures

The estimate of the mean effect size across the 72 effects from reading comprehension measures from both sets of studies was 0.45 ( $p < .001$ , 95% CI = 0.34, 0.57), indicating a moderate positive effect of intervention on students' reading comprehension skills. The variance as measured by the  $Q$  statistic was statistically significant ( $Q = 338.40$ ,  $df = 71$ ,  $p < .001$ ). For the 1980–2004 set of studies, 25 effects from reading comprehension measures were reported, resulting in a mean effect size of 0.91 ( $p < .001$ , 95% CI = 0.59, 1.24), indicating a large effect. For the new set of studies, 47 contributed effects from reading comprehension measures, yielding a mean effect size of 0.24 ( $p < .001$ , 95% CI = 0.16, 0.33), indicating a small effect. Once again, the confidence intervals for the two sets of studies do not overlap, indicating that they come from different populations. The result of a comparison of the mean effect sizes for the 1980–2004 and 2005–2011 sets of studies also indicated that the difference in means was statistically significant ( $Q$ -between = 15.49,  $df = 1$ ,  $p < .001$ ). The variance as measured by the  $Q$  statistic was statistically significant for the 1980–2004 ( $Q = 141.07$ ,  $df = 24$ ,  $p < .001$ ) and 2005–2011 ( $Q = 96.62$ ,  $df = 46$ ,  $p < .001$ ) sets.

### Standardized Reading Comprehension Outcome Measures

The estimate of the mean effect size across the 49 effects from standardized reading comprehension measures from both sets of studies was 0.24 ( $p < .001$ , 95% CI = 0.14, 0.34), indicating a small positive effect of intervention on students' reading comprehension skills. The variance as measured by the  $Q$  statistic was statistically significant ( $Q = 146.93$ ,  $df = 48$ ,  $p < .001$ ). For the 1980–2004 set of studies, the mean effect size across the 10 studies that included standardized reading comprehension measures indicated a moderate effect at 0.65 ( $p = .03$ , 95% CI = 0.06, 1.19). The  $Q$  statistic was statistically significant ( $Q = 73.04$ ,  $df = 9$ ,  $p < .001$ ). For the 39 new studies that reported effects for standardized reading comprehension measures, the mean effect size was smaller at 0.19 ( $p = .03$ , 95% CI = 0.11, 0.27). The  $Q$  statistic was statistically significant ( $Q = 66.26$ ,  $df = 38$ ,  $p = .003$ ). For the standardized reading comprehension outcomes, there is some overlap in the confidence intervals between the 1980–2004

**Table 1.** Characteristics of Intervention Studies.

Study	N	Intervention Type	Person Implementing	Hours of Intervention	Grade Level	Student LD Status	Design Type	Study Set
Abbott & Berninger, 1999	20	Word study	Other	16-25	4th-7th	None LD	MT	1980-2004
Alfassi, 1998	75	Comprehension strategy	Teacher	6-15	9th-12th	None LD	TC	1980-2004
Allinder, Dunse, Brunken, & Obermiller-Krollkowski, 2001	49	Fluency	Teacher	NR	6th-8th	Some LD	TC	1980-2004
Anders, Bos, & Filip, 1983	62	Vocabulary	Other	5 or fewer	9th-12th	All LD	TC	1980-2004
Bhat, Griffin, & Sindelar, 2003	40	Word study	Teacher	NR	6th-8th	All LD	TC	1980-2004
Bhattacharya & Ehri, 2004	40	Word study	Researcher	5 or fewer	6th-9th	None LD	MT	1980-2004
Bos & Anders, 1990	30	Multiple components	Researcher	6-15	NR	All LD	MT	1980-2004
Bos, Anders, Filip, & Jaffe, 1989	50	Multiple components	Other	5 or fewer	NR	All LD	MT	1980-2004
Boyle, 1996	30	Comprehension strategy	Researcher	6-15	6th-8th	Some LD	TC	1980-2004
Chan, 1991	20	Comprehension strategy	Teacher	5 or fewer	5th-6th	All LD	TC	1980-2004
Conte & Humphreys, 1989	26	Fluency	Teacher	6-15	NR	None LD	TC	1980-2004
Darch & Gersten, 1986	24	Comprehension strategy	Other	6-15	9th-12th	All LD	MT	1980-2004
DiCecco & Gleason, 2002	24	Comprehension strategy	Other	6-15	6th-8th	All LD	TC	1980-2004
Fuchs, Fuchs, & Kazdan, 1999	102	Multiple components	Teacher	NR	9th-12th	Some LD	TC	1980-2004
Gajria & Salvia, 1992	30	Comprehension strategy	Researcher	NR	6th-9th	All LD	TC	1980-2004
Hasselbring & Goin, 2004	125	Multiple components	NR	NR	6th-8th	None LD	TC	1980-2004
Homan, Klesius, & Hite, 1993	26	Fluency	Teacher	6-15	6th-8th	None LD	MT	1980-2004
Homan, Hoppes, & Xin, 2000	33	Comprehension strategy	Researcher	NR	6th-8th	Some LD	TC	1980-2004
Jitendra, Hoppes, & Xie, 2000	24	Vocabulary	Researcher	26+	9th-12th	All LD	MT	1980-2004
Johnson, Gersten, & Carnine, 1987	20	Multiple components	Other	26+	NR	All LD	TC	1980-2004
Kennedy & Backman, 1993	26	Comprehension strategy	Researcher	NR	6th-8th	Some LD	MT	1980-2004
Klingner & Vaughn, 1996	24	Multiple components	Teacher	NR	6th-8th	Some LD	TC	1980-2004
Mastropieri et al., 2001	32	Vocabulary	Researcher	NR	7-9th	All LD	MT	1980-2004
Mastropieri, Scruggs, Levin, Gaffney, & McLoone, 1985, Study 1	30	Vocabulary	Researcher	NR	7-9th	All LD	MT	1980-2004
Mastropieri et al., 1985, Study 2	60	Vocabulary	Researcher	5 or fewer	6th-8th	All LD	MT	1980-2004
McLoone, Scruggs, Mastropieri, & Zucker, 1986	21	Comprehension strategy	Teacher	6-15	NR	None LD	TC	1980-2004
Moore & Scevak, 1995	31	Fluency	Other	NR	5th-8th	All LD	MT	1980-2004
O'Shea, Sindelar, & O'Shea, 1987	32	Word study	Teacher	16-25	9th-11th	None LD	TC	1980-2004
Penney, 2002	26	Comprehension strategy	Researcher	6-15	NR	All LD	TC	1980-2004
Snider, 1989	64	Vocabulary	Researcher	5 or fewer	6th-8th	All LD	MT	1980-2004
Veit, Scruggs, & Mastropieri, 1986	91	Comprehension strategy	Teacher	6-15	6th-8th	All LD	MT	1980-2004
Wilder & Williams, 2001	93	Comprehension strategy	Teacher	6-15	5th-8th	All LD	MT	1980-2004
Williams, Brown, Silverstein, & deCani, 1994	40	Comprehension strategy	Teacher	6-15	7th-9th	Some LD	MT	2005-2011
Berkeley, Mastropieri, & Scruggs, 2011	46	Comprehension strategy	Other	6-15	6th-8th	NR	TC	2005-2011

(continued)

Table 1. (continued)

Study	N	Intervention Type	Person Implementing	Hours of Intervention	Grade Level	Student LD Status	Design Type	Study Set
Block, Parris, Reed, Whiteley, & Cleveland, 2009 (data were disaggregated for 4th and 6th graders) <sup>a</sup>	140	Comprehension strategy	Teacher	26+	4th–6th	NR	MT	2005–2011
Burns, Hodgson, Parker, & Fremont, 2011	38	Comprehension strategy	Researcher		6th–8th	NR	MT	2005–2011
Calhoon, 2005	38	Multiple components	Teacher	26+	6th–8th	All LD	TC	2005–2011
Calhoon, Sandow, & Hunter, 2010	60	Comprehension strategy	Researcher	26+	6th–8th	Some LD	MT	2005–2011
Cantrell, Alamsi, Carter, Rintamaa, & Madden, 2010 (data were disaggregated for 6th and 9th graders) <sup>a</sup>	665	Comprehension strategy	Researcher	26+	6 <sup>th</sup> , 9 <sup>th</sup>	All LD	TC	2005–2011
Clarke, Snowling, Truelove, & Hulme, 2010	78	Comprehension strategy	Researcher	26+	4th–5th	NR	MT	2005–2011
Diliberto, Beattie, Flowers, & Algozzine, 2009	74	Fluency	Teacher	6–15	6th–8th	Some LD	TC	2005–2011
Faggella-Luby & Wardwell, 2011 (data were disaggregated for 5th and 6th graders) <sup>a</sup>	82	Comprehension strategy	Other	16–25	5th–6th	None LD	MT	2005–2011
Given, Wasserman, Chari, Beattie, & Eden, 2008	65	Multiple components	Researcher	26+	6th–8th	NR	MT	2005–2011
Graves, Duesbery, Pyle, Brandon, & McIntosh, 2011, Study 1	59	Multiple components	Researcher	26+	6th–8th	Some LD	TC	2005–2011
Graves et al., 2011, Study 2	50	Multiple components	Researcher	26+	6th–8th	Some LD	TC	2005–2011
Guthrie et al., 2009	63	Multiple components	Teacher	26+	4th–5th	NR	TC	2005–2011
Harris, Schumaker, & Deshler, 2011	22	Vocabulary	Researcher	6–15	9th–12th	Some LD	MT	2005–2011
A. Kim et al., 2006	34	Comprehension strategy	Teacher	6–15	6th–8th	All LD	TC	2005–2011
J. S. Kim, Samson, Fitzgerald, & Hartry, 2010	264	Multiple components	Teacher	26+	4th–8th	Some LD	TC	2005–2011
Lang et al., 2009 (data were disaggregated for high-risk and low-risk students) <sup>a</sup>	1,197	Multiple components	Teacher	26+	9th	Some LD	MT	2005–2011
Lovett, Lacerenza, De Palma, & Frijters, 2012	351	Multiple components	Teacher	NR	9th–12th	Some LD	TC	2005–2011
Macaruso & Rodman, 2009	42	Multiple components	Teacher	16–25	6th–8th	NR	TC	2005–2011
Manset-Williamson & Nelson, 2005	20	Comprehension strategy	Researcher	16–25	4th–8th	All LD	MT	2005–2011
McCallum et al., 2011	230	Comprehension strategy	Researcher	5 or fewer	9th–12th	NR	MT	2005–2011
Meyer, Wijekumar, & Lin, 2011	43	Comprehension strategy	Other	16–25	4th–5th	NR	MT	2005–2011
Rasinski, Samuels, Hiebert, Petscher, & Feller, 2011 (data were disaggregated by grade for 4th–10th grades) <sup>a</sup>	658	Multiple components	Teacher	16–25	4th–10th	Some LD	TC	2005–2011
Shippen, Houchins, Steventon, & Sartor, 2005	55	Fluency	Teacher	26+	6th–8th	Some LD	MT	2005–2011
Somers et al., 2010	5,595	Multiple components	Teacher	26+	9th	All LD	MT	2005–2011
Spencer & Manis, 2010	59	Fluency	Other	6–15	6th–8th	Some LD	TC	2005–2011

(continued)

**Table 1. (continued)**

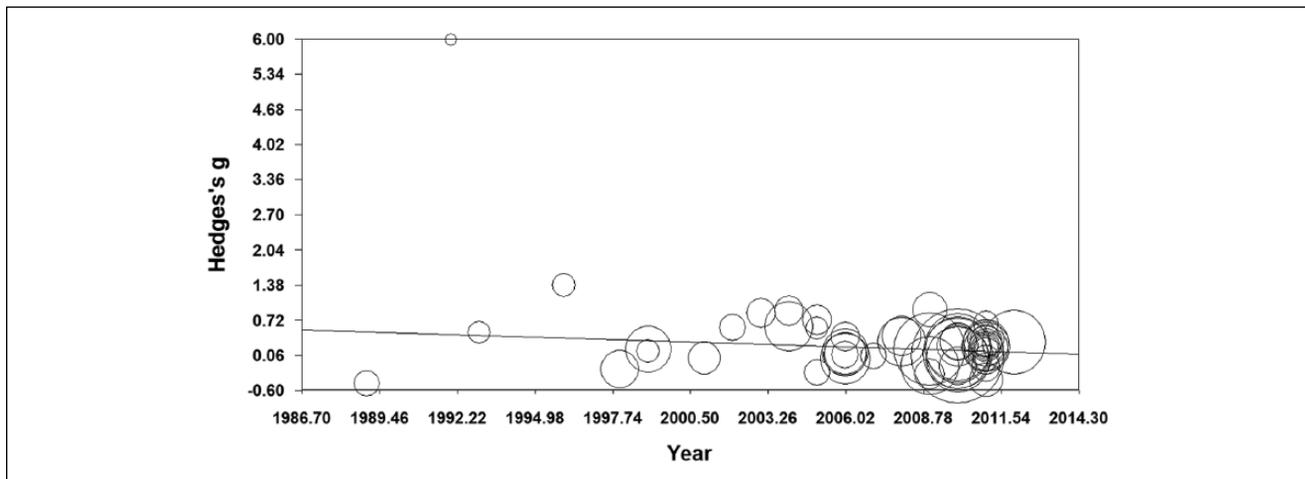
Study	N	Intervention Type	Person Implementing	Hours of Intervention	Grade Level	Student LD Status	Design Type	Study Set
Thames et al., 2008	61	Comprehension strategy	Researcher	26+	4th–8th	NR	TC	2005–2011
Therrien, Wickstrom, & Jones, 2006	29	Multiple components	Other	26+	4th–8th	Some LD	TC	2005–2011
Torgesen et al., 2006, Study 1	126	Multiple components	Teacher	26+	4th–5th	Some LD	TC	2005–2011
Torgesen et al., 2006, Study 2	104	Word study	Teacher	26+	4th–5th	Some LD	TC	2005–2011
Torgesen et al., 2006, Study 3	91	Word study	Teacher	26+	4th–5th	Some LD	TC	2005–2011
Torgesen et al., 2006, Study 4	86	Multiple components	Teacher	26+	4th–5th	Some LD	TC	2005–2011
Vadasy & Sanders, 2008	119	Fluency	Other	16–25	4th–5th	Some LD	TC	2005–2011
Vaughn, Cirino, et al., 2010	327	Multiple components	Researcher	26+	6th	Some LD	TC	2005–2011
Vaughn, Klingner, et al., 2011	95	Comprehension strategy	Teacher	16–25	6th–8th	NR	TC	2005–2011
Vaughn, Wanzek, et al., 2010	476	Multiple components	Researcher	26+	6th–8th	Some LD	MT	2005–2011
Vaughn, Wexler, et al., 2011	182	Multiple components	Researcher	26+	6th–8th	Some LD	MT	2005–2011
Wanzek, Vaughn, Roberts, & Fletcher, 2011	120	Multiple components	Researcher	NR	6th–8th	All LD	TC	2005–2011
Wexler, Vaughn, Roberts, & Denton, 2010	96	Multiple components	Researcher	6–15	9th–12th	Some LD	MT	2005–2011

Note. LD = learning disability; TC = treatment vs. comparison; MT = multiple treatments.

\*In studies where data were disaggregated by grade or by another grouping variable, each grade or group contributed an independent effect size to the meta-analyses presented in this report.



**Figure 1.** Scatterplot of effect size by year of publication for all types of outcome measures.  
Note. Area of the circles on the graph is proportionate to the study's weight.



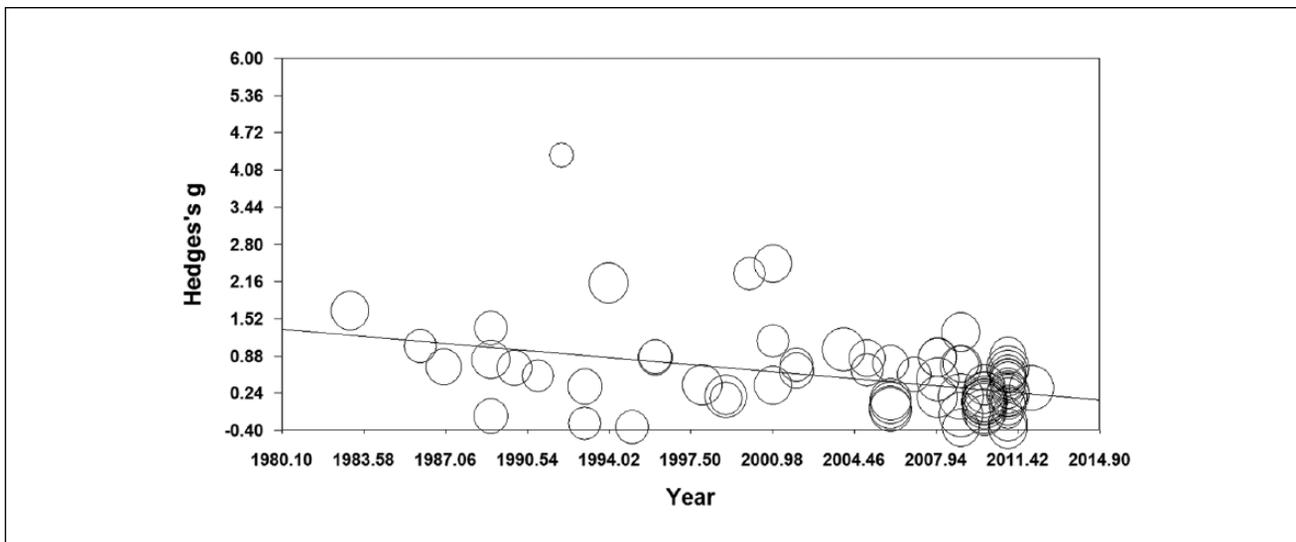
**Figure 2.** Scatterplot of effect size by year of publication for standardized outcome measures.  
Note. Area of the circles on the graph is proportionate to the study's weight.

and 2005–2011 sets of studies. It is possible for statistically significant differences to exist between the mean effect sizes of the two sets of studies with some overlap in the confidence intervals. However, the results of comparison of the mean effect sizes indicated that the difference between the 1980–2004 and 2005–2011 sets of studies was not statistically significant ( $Q$ -between = 2.24,  $df = 1$ ,  $p = .13$ ).

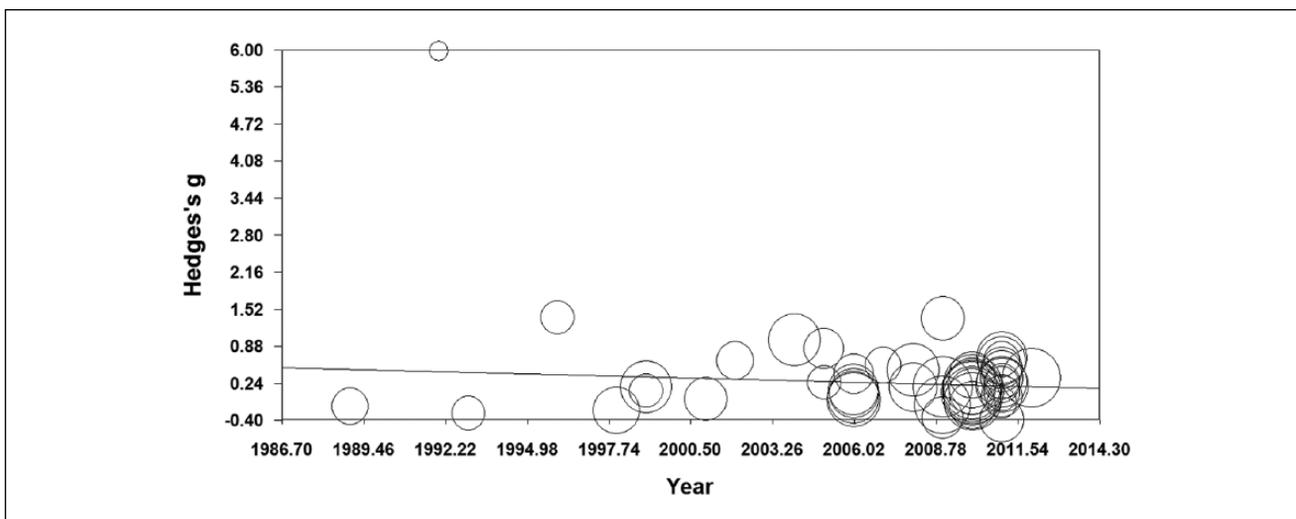
### Changes in Effect Size Magnitude Over Time

The results reported above suggest a decrease in the magnitude of effect sizes for these reading interventions over time. To better understand the relationship between year of publication and effect size, meta-regression was conducted using year of publication as a predictor of effect size in a mixed-effects model using unrestricted maximum

likelihood estimation. Year of publication was a statistically significant predictor of effect size when considering all types of outcome measures ( $\beta = -.04$ ,  $SE = 0.01$ ,  $Q$ -model = 40.95,  $df = 1$ ,  $p < .001$ ,  $T^2 = .15$ ) and all measures of reading comprehension ( $\beta = -.04$ ,  $SE = 0.01$ ,  $Q$ -model = 18.47,  $df = 1$ ,  $p < .001$ ,  $T^2 = .18$ ). However, when only standardized measures of outcomes were included, the mixed effects model indicated that year of publication was not a significant predictor of effect size ( $\beta = -.02$ ,  $SE = 0.01$ ,  $Q$ -model = 3.37,  $df = 1$ ,  $p = .055$ ,  $T^2 = .01$ ). This finding held true for standardized measures of reading comprehension also ( $\beta = -.02$ ,  $SE = 0.01$ ,  $Q$ -model = 1.23,  $df = 1$ ,  $p = .055$ ,  $T^2 = .07$ ). Notably, standardized measures were seldom used in reading intervention studies for Grades 4 through 12 prior to 2000. See Figures 1 to 4 for scatterplots of effect sizes by year of publication.



**Figure 3.** Scatterplot of effect size by year of publication for all measures of reading comprehension. Note. Area of the circles on the graph is proportionate to the study's weight.



**Figure 4.** Scatterplot of effect size by year of publication for standardized measures of reading comprehension. Note. Area of the circles on the graph is proportionate to the study's weight.

The characteristics of studies were examined more closely to determine how changes in study design and participant characteristics over time may have led to reductions in effect sizes. As shown in Table 2, the 1980–2004 group of studies was broken out into smaller spans of years for comparison with the 2005–2011 group of studies and profiled by all moderator variables and two additional research design characteristics that can influence effect sizes and standard errors: number of participants and number of groups. Overall, the 2005–2011 group of studies differs in important

ways from the 1980–2004 group. Studies done from 2005 to 2011 were more likely to have a large number of participants and to include more than two groups. Approximately 78% of the studies done since 2005 provided at least 16 hr of intervention, compared to 18.1% of studies done between 1980 and 2004 and 40% of studies done between 2001 and 2004. Studies published before 2005 also tended to focus exclusively on students with LD, whereas more recent studies were more likely to include a mix of students with LD and struggling readers without an LD designation.

**Table 2.** Comparison of Study Characteristics Over Time.

	1980–1985 (%)	1986–1990 (%)	1991–1995 (%)	1996–2000 (%)	2001–2004 (%)	1980–2004 (%)	2005–2011 (%)
<b>Intervention type</b>							
Comprehension strategy	0.0	22.2	66.7	66.7	25.0	37.5	34.0
Fluency	0.0	22.2	16.7	0.0	12.5	12.5	8.0
Multiple components	0.0	22.2	16.7	16.7	25.0	18.8	52.0
Vocabulary	100.0	33.3	0.0	0.0	0.0	18.8	2.0
Word study	0.0	0.0	0.0	16.7	37.5	12.5	4.0
<b>Grade level</b>							
4th–5th	0.0	0.0	0.0	0.0	0.0	0.0	26.5
6th–8th	0.0	50.0	100.0	60.0	85.7	66.7	52.9
9th–12th	100.0	50.0	0.0	40.0	14.3	33.3	20.6
<b>LD status</b>							
All LD	100.0	100.0	83.3	0.0	37.5	62.5	18.4
Some LD	0.0	0.0	0.0	66.7	37.5	21.9	76.3
None LD	0.0	0.0	16.7	33.3	25.0	15.6	5.3
<b>Type of implementer</b>							
Researcher	100.0	83.3	20.0	60.0	16.7	50.0	39.5
Teacher	0.0	16.7	80.0	40.0	83.3	50.0	60.5
<b>Length of intervention</b>							
5 hr or fewer	100.0	50.0	20.0	0.0	20.0	31.8	4.3
6–15 hr	0.0	50.0	60.0	66.7	40.0	50.0	17.4
16–25 hr	0.0	0.0	0.0	33.3	40.0	13.6	30.4
26+ hr	0.0	0.0	20.0	0.0	0.0	4.5	47.8
<b>Design</b>							
Multiple treatments	66.7	77.8	33.3	33.3	25.0	46.9	40.0
Treatment/comparison	33.3	22.2	66.7	66.7	75.0	53.1	60.0
<b>Number of groups</b>							
2	100.0	88.9	100.0	100.0	77.8	93.8	64.0
3	0.0	0.0	0.0	0.0	11.1	3.1	20.0
4	0.0	11.1	0.0	0.0	0.0	3.1	10.0
5	0.0	0.0	0.0	0.0	0.0	0.0	6.0
<b>Total number of participants</b>							
40 or fewer	66.7	66.7	83.3	66.7	62.5	68.8	17.5
41–60	0.0	22.2	0.0	0.0	12.5	9.4	20.0
61–100	33.3	11.1	16.7	16.7	12.5	15.6	25.0
101–200	0.0	0.0	0.0	16.7	12.5	6.3	15.0
201–400	0.0	0.0	0.0	0.0	0.0	0.0	10.0
401 or more	0.0	0.0	0.0	0.0	0.0	0.0	12.5

Note. LD = learning disability.

### Moderator Analyses

As a result of the lack of overlap in the confidence intervals between the 1980–2004 group of studies and the new (2005–2011) group of studies and the differences in study characteristics across time shown in Table 2, separate moderator analyses were conducted for each set of studies and for the two sets combined.

See Table 3 for *Q*-between statistics and *p* values for each moderator variable analysis. When statistically significant differences were found in moderator analyses, pairwise

comparisons were conducted. To avoid inflating Type I error rates, the *p* value to determine statistical significance was reduced by dividing 0.05 by the number of comparisons made. In some cases, a significant overall *Q*-between statistic yielded no difference in any pairwise comparison at the reduced *p* value selected for the comparisons. In cases where the number of studies at a particular level of a moderator variable was less than four, the studies were not included in the moderator analysis.

See Table 4 for effect sizes, number of studies, and standard errors broken out by each moderator variable. For the

**Table 3.** Results From Moderator Analyses.

Moderator	Study Group	All Outcome Measures			Standardized Outcome Measures			All Reading Comprehension Measures			Standardized Reading Comprehension Measures		
		Q-btwn	df	p	Q-btwn	df	p	Q-btwn	df	p	Q-btwn	df	p
Type of intervention	1980–2004	22.53	4	<.001	10.04	3	.018	8.54	3	.036	4.47	3	.215
	2005–2011	8.95	2	.011	3.26	2	.196	5.93	1	.015 <sup>a</sup>	0.16	1	.692
	Overall	45.29	4	<.001	6.23	3	.101	17.39	3	.001	2.85	3	.415
Grade grouping	1980–2004	1.75	1	.186	2.88	1	.090	0.89	1	.346	1.93	1	.165
	2005–2011	0.73	2	.696	1.55	2	.460	4.35	2	.114	1.81	2	.404
	Overall	3.52	2	.172	3.12	2	.210	3.83	2	.147	3.35	2	.188
Type of implementer	1980–2004	6.26	1	.012	4.76	1	.029	3.11	1	.078	2.39	1	.122
	2005–2011	0.12	1	.729	0.01	1	.916	0.23	1	.630	0.13	1	.719
	Overall	5.97	1	.015	2.25	1	.134	0.25	1	.617	1.43	1	.231
LD status of participants	1980–2004	14.91	2	<.001	1.66	2	.436	8.72	2	.013	2.31	2	.315
	2005–2011	0.12	1	.724	0.14	1	.705	0.42	1	.518	0.00	1	.970
	Overall	21.33	2	<.001	4.02	2	.134	11.83	2	.003	0.66	2	.718
Hours of intervention	1980–2004	2.36	2	.308		ND		0.94	2	.626		ND	
	2005–2011	2.66	2	.265	1.89	2	.389	1.16	2	.561	2.93	2	.231
	Overall	15.78	3	.001	1.92	2	.363	8.67	3	.034 <sup>a</sup>	5.80	2	.055
Design type	1980–2004	1.05	1	.305	0.24	1	.628	0.38	1	.537	0.04	1	.848
	2005–2011	0.00	1	.951	3.53	1	.060	0.03	1	.868	0.39	1	.533
	Overall	3.74	1	.053	0.39	1	.533	1.11	1	.291	0.28	1	.596

Note. LD = learning disability; ND = insufficient data,  $k < 3$ .

<sup>a</sup>Although the moderator analysis indicated that a statistically significant difference was found, pairwise comparisons did not yield a significant difference at the reduced  $p$  value selected to avoid inflating Type I error rates.

standardized outcomes from the new group of studies, the  $Q$  statistic did not indicate that statistically significant variance was present. The effect sizes for this group of studies are reported in Table 3 for each moderator variable for descriptive purposes only.

**Type of intervention.** Interventions were coded based on the type of intervention provided: vocabulary, word study, fluency, comprehension, or multiple components of reading. Nearly all of the multicomponent interventions included instruction targeting fluency and reading comprehension. Two thirds also provided vocabulary instruction. Statistically significant variance as measured by the  $Q$  statistic was found to exist between intervention types. In pairwise comparisons of the 1980–2004 group of studies, comprehension interventions were found to have a significantly greater mean effect size than fluency interventions when all outcome measures were included in the analysis. In addition, vocabulary interventions were found to have a significantly greater mean effect size than fluency, word study, and multicomponent interventions. However, when looking only at standardized outcome measures, the only significant difference was a significantly greater effect for word study interventions over fluency interventions. As would be expected, reading comprehension interventions had a significantly higher mean effect size than fluency interventions when looking only at reading comprehension measures. All differences were statistically significant at  $p < .005$ . In the 2005–2011 group of

studies, pairwise comparisons showed that comprehension interventions had a significantly larger mean effect size than multicomponent interventions when looking at all outcome measures. No other pairwise comparisons differed at  $p < .005$  in the 2005–2011 group of studies.

Results from pairwise comparisons of all studies indicated that vocabulary interventions had a significantly larger mean effect size than all other types of interventions when analyzing all outcome measures. Comprehension interventions were found to have a significantly greater mean effect size than fluency and multicomponent interventions when looking at all outcomes. Again, as would be expected, comprehension interventions were found to have a significantly larger mean effect size on reading comprehension outcomes than fluency, word study, and multicomponent interventions. No other differences were found to be significant at  $p < .005$ .

**LD status.** LD status was coded based on researchers' descriptions of participants. It was seldom the case that researchers described how students in their studies came to be designated as having an LD. Most often, information about LD status was coded from a table of participant demographics in which the number of students with LD was provided or from a brief mention in the participants section of the study that stated that students with LD were included in the intervention or were the focus of the intervention. Therefore, it may be that differences existed

**Table 4.** Statistics by Moderator Variables.

Moderator	Level	Study Group	All Outcome Measures			Standardized Outcome Measures			All Reading Comprehension Measures			Standardized Reading Comprehension Measures			
			ES	k	SE	ES	k	SE	ES	k	SE	ES	k	SE	
Type of intervention	Reading comprehension	1980–2004	1.23	12	0.28	2.25 <sup>a</sup>	3	1.34	1.34	12	0.30	2.24 <sup>a</sup>	3	1.35	
		2005–2011	0.40	17	0.08	0.21	10	0.06	0.39	17	0.07	0.21	10	0.06	
		Overall	0.74	29	0.12	0.47	13	0.15	0.78	29	0.13	0.46	13	0.15	
	Fluency	1980–2004	0.24 <sup>a</sup>	4	0.28	ND			0.33 <sup>a</sup>	4	0.18	ND			
		2005–2011	0.31	4	0.13	0.28	4	0.12	ND			ND			
		Overall	0.30	8	0.10	0.17 <sup>a</sup>	6	0.13	0.31	7	0.13	0.21 <sup>a</sup>	5	0.11	
	Word study	1980–2004	0.60	4	0.18	0.68	4	0.18	ND			ND			
		2005–2011	ND			ND			ND			ND			
		Overall	0.33	6	0.13	0.39	6	0.16	0.13 <sup>a</sup>	4	0.13	0.13 <sup>a</sup>	4	0.13	
	Vocabulary	1980–2004	1.59	6	0.27	ND			ND			ND			
		2005–2011	ND			ND			ND			ND			
		Overall	1.58	7	0.24	ND			ND			ND			
	Multiple components	1980–2004	0.55	6	0.13	0.41	3	0.15	0.60	6	0.21	0.36 <sup>a</sup>	3	0.36	
		2005–2011	0.14	26	0.05	0.11	25	0.04	0.17	25	0.06	0.18	24	0.06	
		Overall	0.20	32	0.05	0.14	28	0.04	0.24	31	0.06	0.46	27	0.06	
Grade grouping	4th–5th	1980–2004	ND			ND			ND			ND			
		2005–2011	0.30	9	0.10	0.22	8	0.09	0.30	9	0.13	0.29	8	0.14	
		Overall	0.30	9	0.10	0.22	8	0.09	0.30	9	0.13	0.29	8	0.14	
	6th–8th	1980–2004	1.10	12	0.20	0.65	4	0.24	1.11	9	0.25	0.75 <sup>a</sup>	3	0.39	
		2005–2011	0.25	18	0.08	0.16	14	0.06	0.29	17	0.09	0.20	13	0.09	
		Overall	0.57	30	0.10	0.25	18	0.08	0.55	26	0.12	0.30	16	0.10	
	9th–12th	1980–2004	0.67	6	0.26	0.13 <sup>a</sup>	3	0.19	0.75	5	0.29	0.14 <sup>a</sup>	3	0.21	
		2005–2011	0.18 <sup>a</sup>	7	0.11	0.09 <sup>a</sup>	5	0.07	0.10 <sup>a</sup>	6	0.06	0.10 <sup>a</sup>	5	0.07	
		Overall	0.35	13	0.11	0.09 <sup>a</sup>	8	0.06	0.28	11	0.10	0.10 <sup>a</sup>	8	0.06	
	Type of implementer	Researcher	1980–2004	1.48	12	0.25	2.59	3	1.11	1.61	6	0.43	ND		
			2005–2011	0.19	17	0.06	0.14	12	0.05	0.20	16	0.07	0.16	12	0.07
			Overall	0.68	29	0.11	0.33	15	0.12	0.46	22	0.11	0.34	14	0.13
		Teacher	1980–2004	0.63	12	0.22	0.15 <sup>a</sup>	6	0.18	0.71	11	0.28	0.06 <sup>a</sup>	5	0.13
			2005–2011	0.22	26	0.06	0.13	23	0.05	0.25	25	0.06	0.19	22	0.06
			Overall	0.35	38	0.07	0.13	29	0.05	0.40	36	0.08	0.17	27	0.05
LD status of participants	All designated learning disabled	1980–2004	1.20	20	0.18	1.55 <sup>a</sup>	4	0.90	1.13	14	0.26	1.73 <sup>a</sup>	3	1.43	
		2005–2011	0.19	7	0.08	0.15	7	0.06	0.20	7	0.08	0.14	7	0.05	
		Overall	0.95	27	0.14	0.41	11	0.14	0.83	21	0.15	0.33	10	0.14	
	Some designated learning disabled, some struggling	1980–2004	0.79	7	0.23	0.47	4	0.23	0.89	7	0.23	0.59	4	0.30	
		2005–2011	0.16	29	0.05	0.12	26	0.04	0.15	26	0.05	0.15	24	0.05	
		Overall	0.32	7	0.13	0.15	30	0.04	0.27	33	0.07	0.20	28	0.06	
	All struggling, none designated learning disabled	1980–2004	0.28 <sup>a</sup>	5	0.16	0.33 <sup>a</sup>	4	0.27	0.27 <sup>a</sup>	4	0.19	0.12 <sup>a</sup>	3	0.27	
		2005–2011	ND			ND			ND			ND			
		Overall	0.24	36	0.06	0.35 <sup>a</sup>	6	0.19	0.32	6	0.15	0.21 <sup>a</sup>	5	0.20	
Hours of intervention	0–5 hr	1980–2004	1.18	7	0.26	ND			1.05	3	0.34	ND			
		2005–2011	ND			ND			ND			ND			
		Overall	1.00	9	0.27	ND			0.78	5	0.32	ND			
	6–15 hr	1980–2004	0.79	11	0.24	ND			0.88	11	0.28	ND			
		2005–2011	0.44	8	0.15	0.22	5	0.13	0.34	6	0.16	0.10	4	0.13	
		Overall	0.66	19	0.16	0.08	7	0.12	0.69	17	0.20	0.01	6	0.11	
	16–25 hr	1980–2004	0.62	3	0.27	ND			0.64	3	0.27	ND			
		2005–2011	0.23	14	0.08	0.21	14	0.08	0.32	14	0.08	0.32	14	0.08	
		Overall	0.27	17	0.07	0.22	16	0.07	0.35	17	0.08	0.32	16	0.08	
	26+ hr	1980–2004	ND			ND			ND			ND			
		2005–2011	0.18	22	0.06	0.10	18	0.05	0.22	21	0.07	0.17	17	0.07	
		Overall	0.18	23	0.06	0.10	19	0.05	0.21	22	0.07	0.16	18	0.06	

(continued)

Table 4. (continued)

Moderator	Level	Study Group	All Outcome Measures			Standardized Outcome Measures			All Reading Comprehension Measures			Standardized Reading Comprehension Measures			
			ES	k	SE	ES	k	SE	ES	k	SE	ES	k	SE	
Design type	Multiple treatment	1980–2004	1.10	15	0.19	0.81	3	0.33	1.05	9	0.28	ND			
		2005–2011	0.23	20	0.07	0.07	14	0.03	0.25	18	0.07	0.15	13	0.07	
		Overall	0.62	35	0.10	0.17	17	0.07	0.54	27	0.10	0.20	15	0.07	
	Treatment vs. control	1980–2004	0.83	17	0.18	0.60	9	0.28	0.83	16	0.20	0.61 <sup>a</sup>	8	0.33	
		2005–2011	0.23	30	0.06	0.17	27	0.04	0.24	29	0.06	0.21	26	0.06	
		Overall	0.39	47	0.07	0.22	36	0.06	0.41	45	0.07	0.25	34	0.07	

Note. ND = insufficient data,  $k < 3$ .

<sup>a</sup>95% confidence interval includes zero.

between studies in the way LD status was determined. In studies where both struggling readers without LD and struggling readers with LD participated, both were randomly assigned to treatment without regard for their LD status. In these studies, no researchers reported differences in the alternative instruction provided to participants with LD and without LD in the comparison group.

In pairwise comparisons, a statistically significant difference in mean effect size was found between studies that included only students with LD and those that included only struggling readers who were not designated as having an LD. This difference, which favored studies with only students with LD, was found to be significant at  $p < .017$  in the 1980–2004 group of studies when including all outcomes and only reading comprehension outcomes. When looking across all 82 studies, the mean effect size for studies that included only students with LD was significantly greater at  $p < .017$  than the mean effect size for studies that included some students with LD and some struggling readers and studies that did not include any students with LD. This finding held true when analyzing all outcomes and only reading comprehension outcomes. No significant differences were found based on the LD status of participants in the 2005–2011 group of studies.

**Hours of intervention.** The number of hours of intervention provided was examined as a categorical variable because many studies reported this information as a range of hours rather than a single number. Treating hours of intervention as a continuous variable would have resulted in missing values for these studies, excluding them from the analyses. In pairwise comparisons, shorter interventions were found to have a significantly larger mean effect size when analyzing all outcomes across all studies, with studies that provided 5 hr of intervention or less and those that provided 6 to 15 hr of intervention having larger mean effects than studies that provided 26 hr of intervention or more. These differences were statistically significant at  $p < .008$ . No differences

based on number of hours of intervention were found to be statistically significant in pairwise comparisons at  $p < .008$  when looking at the 1980–2004 and 2005–2011 groups of studies separately.

**Type of implementer.** Studies in which researchers implemented the intervention had a significantly larger mean effect size than studies in which teachers implemented the intervention when looking at all outcome measures across all studies at  $p < .05$ . This finding held true when analyzing only the 1980–2004 group of studies for all types of outcome measures and standardized measures only. No differences based on who implemented the intervention were statistically significant when considering the 2005–2011 group of studies only.

**Grade level.** No significant differences were found in pairwise comparisons between mean effect sizes for studies that included only 4th–5th graders, only 6th–8th graders, and only 9th–12th graders. This finding held true for the 1980–2004 and 2005–2011 sets of studies separately and for all studies combined. However, in the 2005–2011 set of studies the confidence intervals for the mean effect sizes for students in Grades 9 to 12 includes zero, meaning that it is possible (though not likely) that reading interventions have no effect on high school students.

**Design type.** No significant differences were found between mean effect sizes for studies that compared multiple treatments and those that compared a treatment and control or comparison condition. This finding held true for the 1980–2004 and 2005–2011 sets of studies separately and for all studies combined.

### Publication Bias

Publication bias was evaluated using the trim-and-fill approach. This approach uses an iterative technique, removing studies

**Table 5.** Publication Bias Analysis.

		All Outcome Measures	Standardized Outcome Measures	All Reading Comprehension Measures	Standardized Reading Comprehension Measures
Number of studies missing	1980–2004	4	4	5	3
	2005–2011	7	7	12	8
	Overall	11	0	10	5
Adjusted effect size and 95% confidence interval	1980–2004	1.10 (0.82, 1.37)	1.11 (0.58, 1.64)	1.17 (0.83, 1.50)	1.06 (0.44, 1.68)
	2005–2011	0.17 (0.09, 0.26)	0.11 (0.04, 0.17)	0.12 (0.02, 0.21)	0.10 (0.01, 0.20)
	Overall	0.61 (0.49, 0.72)	—	0.57 (0.45, 0.69)	0.29 (0.19, 0.39)

causing a lack of symmetry in the funnel plot of effect sizes and calculating a mean effect, and then returning these studies and adding in imputed effects to create a symmetrical plot (Card, 2012). The purpose of the analysis is to determine if estimates of mean effect size were biased by the exclusion of effect sizes from nonpublished research and published studies that might have been missed in the literature search. Results indicated that publication bias did not affect the mean effect size estimates for the meta-analysis of standardized outcome measures for all 82 studies. For all of the other meta-analyses conducted for this report, the trim-and-fill analyses found some evidence of publication bias. See Table 5 for the number of studies that are estimated to be missing from each meta-analysis and the adjusted mean effect size estimates and 95% confidence intervals that result from including imputed values for missing studies. Imputing the effect sizes for the missing studies did not result in any 95% confidence intervals that included zero.

## Discussion

This meta-analysis of reading interventions conducted between 1980 and 2011 for students in Grades 4 through 12 with reading difficulties was intended as an update to and extension of Scammacca et al. (2007), adding studies published since 2004 and drawing conclusions about new learning on the effectiveness of interventions for struggling readers in Grades 4 to 12. Because the more recent studies were representative of a different population of studies than those analyzed in Scammacca et al., we conducted separate meta-analyses for each set of studies to compare the findings across the two sets. Meta-analyses also were conducted with the two groups combined to see what conclusions could be drawn based on the full corpus of research.

### *Effectiveness of Reading Interventions for Students in Grades 4–12*

Based on the results of the meta-analyses, it is clear that reading interventions produce positive results for students in Grades 4 to 12. Across all 30 years of studies and

including all reading outcome measures, the benefit of intervention was an increase of nearly one half of one standard deviation. Results from standardized measures indicated that the gains were somewhat smaller, around one fifth of one standard deviation for students receiving intervention. On measures of reading comprehension, results also showed that students benefitted from intervention, with effect sizes of similar magnitude to those found when considering all types of reading measures. A decline in effect sizes over time was observed. Studies conducted between 1980 and 2004 resulted in larger effects than those conducted between 2005 and 2011.

When interpreting the effect size for reading interventions, it is important to compare them to typical yearly gains in reading ability for students in these grades. Bloom, Hill, Black, and Lipsey (2008) computed average gains by students over one academic year on seven nationally normed measures of reading achievement. They report annual growth effect sizes ranging from 0.40 for students in Grade 4 to 0.06 for students in Grade 11, with effects decreasing in a linear fashion as grade level increases. The effect sizes found for reading interventions in the present meta-analyses compare favorably to these annual growth effect sizes. These interventions typically lasted less than a full academic year, yet produced effects that on average were close to one year's growth when all types of reading measures are considered. In judging the effect sizes for standardized measures, it is important to note that scores on these measures are based on norms that take expected academic growth into account. Therefore, the smaller effect sizes observed for these measures represent gains in addition to what would be expected due to typical instruction and developmental growth. Based on this information, it appears that reading interventions make a positive difference for struggling readers in Grades 4 to 12.

The only statistically significant moderator of effectiveness for the 2005–2011 group of studies was type of intervention. Reading comprehension interventions were associated with significantly higher effect sizes than fluency interventions in the meta-analysis of all types of outcomes. This finding was also true for the 1980–2004 group

of studies, indicating that the difference in favor of reading comprehension interventions is somewhat robust despite the fact that it was not observed when considering standardized measures in either group of studies. Vocabulary interventions, which were associated with very large mean effect sizes that were significantly higher than other intervention types in the 1980–2004 group of studies, were mostly absent from the 2005–2011 group of studies and are rarely evaluated using standardized measures (Scammacca et al., 2007). Two thirds of the multiple-component interventions in the 2005–2011 group of studies included some vocabulary instruction. It may be that the large effect sizes reported in previous meta-analyses persuaded researchers to include a vocabulary component in their interventions.

### *Accounting for Differences Across Time*

Findings indicated that more recent studies yielded substantially smaller mean effect sizes than the older studies. Results of meta-regression analyses indicated that year of publication predicted effect size when analyzing effect sizes from all types of outcome measures and all types of reading comprehension measures, but not when effect sizes from standardized measures were analyzed. These results indicate that the increased use of standardized measures in more recent studies is one important factor in the decrease seen in effect size over time. This interpretation is based on consistent reporting that the use of standardized measures in intervention research is associated with smaller effect sizes (Swanson, Hoskyn, & Lee, 1999; Willingham, 2007). Willingham (2007) suggests that experimenter-designed measures of reading comprehension tend to use reading passages that are amenable to the strategies that were taught in the intervention, whereas standardized measures use a variety of passages that may require students to apply strategies not taught in the intervention or to apply strategies that were taught in new ways.

Another possible cause of shifts in effect size over time is the changing nature of the instruction provided to the comparison group. When a study compares multiple treatments or compares treatment to a business-as-usual comparison group that is receiving an intervention provided by the school, the study-wise effect size reflects the added benefit of one intervention over another, not the benefit of intervention over no intervention. It was difficult to nail down the exact nature of the instruction provided in business-as-usual comparison groups in most of the studies that used them because the research reports tended not to describe the comparison group's instruction in sufficient detail. We attempted to evaluate the business-as-usual conditions in the 1980–2004 and 2005–2011 groups of studies by comparing scores on standardized measures in each group. This effort proved fruitless due to differences in how scores were reported, differences in measures used, and

differences in forms between older and current versions of measures. Nevertheless, the similarity in effect sizes between studies that used a multiple-treatment design and those that used a treatment-comparison design suggests that the comparison conditions likely involved some type of instruction. With the increasing implementation of RTI models, these comparison-group interventions may be more intensive than in the past.

The 1980–2004 and 2005–2011 groups of studies also differ in the populations of students who participated in them. Identification of students in need of intervention based in part on RTI criteria has led to fewer studies that focus exclusively on students with LD and more that include struggling readers who have not been classified as having LD. In more recent studies, students who are designated as having LD based on criteria other than the IQ–achievement discrepancy may have lower IQs than students with LD who were included in earlier studies based on IQ–achievement discrepancies. In addition, a larger percentage of the more recent studies used teachers to deliver the intervention. Teachers were shown to be effective at delivering interventions in Scammacca et al. (2007), but researcher-led interventions had significantly larger effects. The studies that have been conducted more recently depend on teachers to deliver the interventions because lengthy interventions and large sample sizes make it cost-prohibitive to employ researchers to deliver the interventions.

Another key difference between the 1980–2004 and 2005–2011 groups of studies is the length of the interventions. More than three out of four of the studies published between 2005 and 2011 provided at least 16 hr of intervention, compared to less than 20% of studies published between 1980 and 2004. On the surface, it seems counterintuitive to state that longer interventions are associated with smaller effects. A possible explanation for the negative relationship found between effect size and hours of intervention is posited by Willingham (2007, 2012). He claimed that brief reading comprehension interventions (5 hr or less) can produce a large immediate effect for students who are adequate decoders because reading comprehension strategies are easy to learn. He asserted that maintaining the gain seen in a brief intervention requires that students remember to use the strategies over a longer period of time with new texts that are not similar to the passages used to practice the strategies (such as those on standardized measures of reading comprehension). Although not a completely satisfying explanation for the phenomenon noted in the present meta-analysis, Willingham's theory provides an avenue for future research on the relative effects of brief and extensive interventions.

It is important to note that other features of more recent interventions may be confounded with the length of intervention and explain the reduction in observed effect sizes over time. These features include more precise measurement,

research designs that compared multiple groups, the use of multiple indicators of effectiveness that included proximal and distal measures, increased implementation of randomized controlled trials, changes in the types of students targeted for intervention, and improvements to business-as-usual instruction provided to comparison groups, among others. The finding that longer interventions were associated with smaller effect sizes should not be taken to mean that briefer interventions are more beneficial to students than more extensive interventions. Rather, additional research is needed that measures students' progress at multiple points along the course of a long intervention to determine how estimates of the effect of intervention change over time.

### *Comparisons With Other Recent Meta-Analyses*

The size of the effects observed in the more recent studies is in line with those found in other meta-analyses of reading interventions for students in Grades 4 to 12 that have been published recently. The 12 studies in Flynn et al. (2012) used standardized measures and yielded a mean effect size of 0.41, nearly identical to the mean effect size of 0.42 reported for standardized measures in the 1980–2004 group of studies but larger than 0.13 mean effect size for the 2005–2011 group of studies. Flynn et al. included only students in Grades 5 to 9 with a reading disability. The inclusion of interventions for students in 10th–12th grade and struggling readers who were not identified as having a reading disability in the present meta-analyses might account in part for the lower mean effect sizes found for standardized measures.

In their meta-analysis on reading interventions that provided at least 75 sessions to students in Grades 4 through 12, Wanzek et al. (2013) found mean effect sizes ranged from 0.10 (reading comprehension outcomes) to 0.16 (reading fluency and word-reading fluency outcomes), comparable to the results of the present meta-analyses. For interventions lasting 26 hr or more, the mean effect size found in the 2005–2011 group of studies ranged from 0.10 for standardized measures to 0.22 for all reading comprehension measures. In the analysis of hours of intervention for all studies combined, shorter interventions of 15 hr or less were found to have significantly larger mean effect sizes than interventions that provided 26 hr or more of intervention when outcomes from all measures and all reading comprehension measures were included. These findings support Wanzek et al.'s assertion that longer interventions are associated with smaller effect sizes, though as discussed above the cause of this phenomenon is unclear and further research is needed on the growth curves of students in lengthy interventions.

Flynn et al. (2012) and Wanzek et al. (2013) were unable to find moderator variables that explained the variability present in the effect sizes in their meta-analyses. Similarly,

the 2005–2011 set of studies included in the present meta-analyses exhibited less systematic variation than the 1980–2004 group of studies. Where statistically significant variability was present in the more recent studies, it could not be attributed to the research design, differences in the length of the intervention, the grade level or LD status of the participants, or whether the intervention was implemented by teachers or researchers. This finding is due at least in part to the use of a random effects model.

### *Implications for Practice Based on the Most Current Research*

The current research base supports providing interventions at both the word and text level. Most of the more recent studies focused on reading comprehension (17 studies) or included multiple components (26 studies). The multiple-component interventions included both word-level and text-level instruction and produced a small but statistically significant mean effect. Vocabulary instruction, which was found to be highly effective in the 1980–2004 group of studies, was integrated into two thirds of the multiple component interventions, indicating that researchers may have taken the Scammacca et al. (2007) finding into account when designing their interventions.

In the 2005–2011 group of studies, the mean effect size for reading comprehension interventions reflected an average gain of nearly half a standard deviation when looking at all measures and about one fourth of a standard deviation on standardized measures. Reading comprehension instruction was included in nearly all of the multicomponent interventions. The research base continues to show that teaching reading comprehension strategies to struggling readers in Grades 4 to 12 is beneficial.

In addition, the most current research affirms that teachers can provide effective reading interventions. The mean effect sizes for teacher- and researcher-provided interventions in the 2005–2011 group of studies were nearly identical, both on all measures and on standardized measures. A greater proportion of the studies in the 2005–2011 group used teachers to implement the intervention (26 of 50, 52.0%, compared to 12 of 32, 37.5%, in the 1980–2004 group of studies). The largest and most rigorous studies relied on teachers to implement the intervention (e.g., Lang et al., 2009, with  $N = 1,197$ ; Somers et al., 2010, with  $N = 5,595$ ). Therefore, it appears that teachers increasingly are being trained as interventionists and are proving to be as effective as researchers at providing interventions.

Finally, the most current data show that reading interventions are effective both for struggling readers with LD and those not identified as having LD. No differences based on LD status were found when looking only at the 2005–2011 group of studies. Notably, most of the studies in the 2005–2011 group included both students with and without

LD. Therefore, the most recent research suggests that all struggling readers benefit from intervention regardless of their diagnosed LD status.

### Limitations

The findings presented in this report are limited by the available research literature on interventions for struggling readers in Grades 4 to 12. Some research reports failed to give sufficient detail to allow for coding on all moderator variables of interest. As a result, pairwise comparisons could not be conducted for all levels of all moderators to determine the relative effectiveness of all attributes of interest. In addition, hours of intervention had to be coded as a categorical variable because exact counts of hours were not provided in many of the studies. Coding hours as a continuous variable would have yielded more precise information on the relationship between length of intervention and effectiveness. The counterintuitive finding of larger effects from shorter interventions remains unexplained. Finally, the meta-analyses presented here are limited by the finding that the 1980–2004 and 2005–2011 groups of studies came from different populations of studies. The original analysis plan that called for all studies between 1980 and 2011 to be meta-analyzed together had to be modified to include separate meta-analyses for each group. As a result, the statistical power within the moderator analyses was lower for the separate meta-analyses of the 1980–2004 and 2005–2011 groups of studies than it was when all studies were combined.

### Conclusions

The results presented in this report support the efficacy of reading interventions for struggling readers in Grades 4 to 12, though the magnitude of the effects obtained may be less than originally thought based on the results of Scammacca et al. (2007). More recent research on these interventions have included more rigorous measures of results that capture the extent to which the skills gained through the interventions generalize beyond the immediate context of the intervention. In addition, studies are providing more hours of intervention and increasingly these interventions are compared to an alternative intervention instead of a true no-intervention control group. As a result, smaller effects are observed. Despite these smaller effects, the more recently published interventions likely are more representative of the kind of intervention that struggling readers need. Reading difficulties that have perseverated past the primary school years likely do require many hours of intervention to remediate. Progress is likely to be slow but steady. Teachers are better positioned than researchers to provide longer-term interventions. The good news is that strong evidence exists showing that students in Grades 4 through 12 who are

struggling in reading can improve when targeted with appropriate interventions.

Future research on reading interventions for students in Grades 4 to 12 is needed to confirm and extend what has been learned to date. The nature of the relationship between effect size and length of intervention must be clarified so that educators can make the best use of their time by providing the appropriate dosage of intervention to struggling readers to produce a meaningful and long-term effect. Additional research also might be directed toward improving the knowledge base concerning component skills so that multiple-component interventions can be more effective. For example, if more is known about how to teach word study effectively to older students with reading difficulties, better multiple-component interventions can be designed to capitalize on these findings. Further research also is needed on the current state of business-as-usual interventions for struggling readers. Gaining a better understanding of the nature of what schools are currently providing will allow researchers to craft studies that examine ways to maximize gains for struggling readers in Grades 4 to 12.

### 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.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by Grant P50 HD052117 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development and by the Institute of Education Sciences, U.S. Department of Education, through Grant R305F100013 to the University of Texas at Austin as part of the Reading for Understanding Research Initiative. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health and Human Development, the National Institutes of Health, the Institute of Education Sciences, or the U.S. Department of Education.

### References

- References marked with an asterisk indicate studies included in the meta-analyses.
- \*Abbott, S. P., & Berninger, V. W. (1999). It's never too late to remediate: Teaching word recognition to students with reading disabilities in grades 4–7. *Annals of Dyslexia*, 49, 223–250.
  - \*Alfassi, M. (1998). Reading for meaning: The efficacy of reciprocal teaching in fostering reading comprehension in high school students in remedial reading classes. *American Educational Research Journal*, 35, 309–332.
  - \*Allinder, R. M., Dunse, L., Brunken, C. D., & Obermiller-Krolikowski, H. J. (2001). Improving fluency in at-risk readers

- and students with learning disabilities. *Remedial and Special Education*, 22, 48–45.
- \*Anders, P. L., Bos, C. S., & Filip, D. (1984). The effect of semantic feature analysis on the reading comprehension of learning-disabled students. In J. S. Niles & L. A. Harris (Eds.), *Changing perspectives on reading/language processing and instruction* (pp. 162–166). Rochester, NY: National Reading Conference.
- \*Berkeley, S., Mastropieri, M. A., & Scruggs, T. E. (2011). Reading comprehension strategy instruction and attribution retraining for secondary students with learning and other mild disabilities. *Journal of Learning Disabilities*, 44, 18–32. doi:10.1177/0022219410371677
- \*Bhat, P., Griffin, C. C., & Sindelar, P. T. (2003). Phonological awareness instruction for middle school students with learning disabilities. *Learning Disability Quarterly*, 26, 73–87.
- \*Bhattacharya, A., & Ehri, L. C. (2004). Graphosyllabic analysis helps adolescent struggling readers read and spell words. *Journal of Learning Disabilities*, 37, 331–348.
- \*Biggs, M. C., Homan, S. P., Dedrick, R., Minick, V., & Rasinski, T. (2008). Using an interactive singing software program: A comparative study of struggling middle school readers. *Reading Psychology*, 29, 195–213. doi:10.1080/02702710802073438
- \*Block, C., Parris, S. R., Reed, K. L., Whiteley, C. S., & Cleveland, M. D. (2009). Instructional approaches that significantly increase reading comprehension. *Journal of Educational Psychology*, 101, 262–281. doi:10.1037/a0014319
- Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Educational Effectiveness*, 1, 289–328. doi:10.1080/1934574080240072
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2011). *Comprehensive meta analysis* (Version 2.2.064). Englewood, NJ: Biostat.
- \*Bos, C. S., & Anders, P. L. (1990). Effects of interactive vocabulary instruction on the vocabulary learning and reading comprehension of junior-high learning disabled students. *Learning Disability Quarterly*, 13, 31–42.
- \*Bos, C. S., Anders, P. L., Filip, D., & Jaffe, L. E. (1989). The effects of an interactive instructional strategy for enhancing reading comprehension and content area learning for students with learning disabilities. *Journal of Learning Disabilities*, 22, 384–390.
- \*Boyle, J. R. (1996). The effects of a cognitive mapping strategy on the literal and inferential comprehension of students with mild disabilities. *Learning Disabilities Quarterly*, 19, 86–98.
- \*Burns, M. K., Hodgson, J., Parker, D. C., & Fremont, K. (2011). Comparison of the effectiveness and efficiency of text previewing and preteaching keywords as small-group reading comprehension strategies with middle-school students. *Literacy Research and Instruction*, 50, 241–252. doi:10.1080/19388071.2010.519097
- \*Calhoun, M. (2005). Effects of a peer-mediated phonological skill and reading comprehension program on reading skill acquisition for middle school students with reading disabilities. *Journal of Learning Disabilities*, 38, 424–433. doi:10.1177/00222194050380050501
- \*Calhoun, M., Sandow, A., & Hunter, C. V. (2010). Reorganizing the instructional reading components: Could there be a better way to design remedial reading programs to maximize middle school students with reading disabilities' response to treatment? *Annals of Dyslexia*, 60, 57–85. doi:10.1007/s11881-009-0033-x
- \*Cantrell, S., Alamsi, J., Carter, J., Rintamaa, M., & Madden, A. (2010). The impact of a strategy-based intervention on the comprehension and strategy use of struggling adolescent readers. *Journal of Educational Psychology*, 102, 257–280. doi:10.1037/a0018212
- Card, N. (2012). *Applied meta-analysis for social science research*. New York, NY: Guilford.
- \*Chan, K. S. (1991). Promoting strategy instruction generalization through self-instructional training in students with reading disabilities. *Journal of Learning Disabilities*, 24, 427–290.
- Cheung, M. W. L. (2008). A model for integrating fixed-, random-, and mixed-effects meta-analyses in structural equation modeling. *Psychological Methods*, 13, 182–202. doi:10.1037/a0013163
- \*Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science*, 21, 1106–1116. doi:10.1177/0956797610375449
- \*Conte, R., & Humphreys, R. (1989). Repeated readings using audiotaped material enhances oral reading in children with reading difficulties. *Journal of Communication Disorders*, 22, 65–79.
- \*Darch, C., & Gersten, R. (1986). Direct setting activities in reading comprehension: A comparison of two approaches. *Learning Disability Quarterly*, 9, 235–243.
- \*DiCecco, V. M., & Gleason, M. M. (2002). Using graphic organizers to attain relational knowledge from expository texts. *Journal of Learning Disabilities*, 35, 306–320.
- \*Diliberto, J. A., Beattie, J. R., Flowers, C. P., & Algozzine, R. F. (2009). Effects of teaching syllable skills instruction on reading achievement in struggling middle school readers. *Literacy Research and Instruction*, 48, 14–27. doi:10.1080/19388070802226253
- Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C. K., Cable, A., & Tackett, K. K. (2009). A synthesis of reading interventions and effects on reading outcomes for older struggling readers. *Review of Educational Research*, 79, 262–300. doi:10.3102/0034654308325998
- \*Faggella-Luby, M., & Wardwell, M. (2011). RTI in a middle school: Findings and practical implications of a tier 2 reading comprehension study. *Learning Disability Quarterly*, 34, 35–49.
- Flynn, L. J., Zheng, X., & Swanson, H. (2012). Instructing struggling older readers: A selective meta-analysis of intervention research. *Learning Disabilities Research & Practice*, 27, 21–32. doi:10.1111/j.1540-5826.2011.00347.x
- \*Fuchs, L. S., Fuchs, D., & Kazdan, S. (1999). Effects of peer-assisted learning strategies on high school students with serious reading problems. *Remedial and Special Education*, 20, 309–319.

- \*Gajria, M., & Salvia, J. (1992). The effects of summarization instruction on text comprehension of students with learning disabilities. *Exceptional Children, 58*, 508–516.
- \*Given, B. K., Wasserman, J. D., Chari, S. A., Beattie, K., & Eden, G. F. (2008). A randomized, controlled study of computer-based intervention in middle school struggling readers. *Brain and Language, 106*, 83–97. doi:10.1016/j.bandl.2007.12.001
- \*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. *Elementary School Journal, 111*, 641–661. doi:10.1086/659036
- \*Guthrie, J. T., McRae, A., Coddington, C. S., Klauda, S., Wigfield, A., & Barbosa, P. (2009). Impacts of comprehensive reading instruction on diverse outcomes of low- and high-achieving readers. *Journal of Learning Disabilities, 42*, 195–214. doi:10.1177/0022219408331039
- \*Harris, M., Schumaker, J., & Deshler, D. (2011). The effects of strategic morphological analysis instruction on the vocabulary performance of secondary students with and without disabilities. *Learning Disability Quarterly, 34*, 17–33.
- \*Hasselbring, T. S., & Goin, L. I. (2004). Literacy instruction for older struggling readers: What is the role of technology? *Reading and Writing Quarterly, 20*, 123–144.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*, 107–128.
- \*Homan, S. P., Klesius, J. P., & Hite, C. (1993). Effects of repeated readings and nonrepetitive strategies on students' fluency and comprehension. *Journal of Educational Research, 87*, 94–99.
- Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum.
- Individuals with Disabilities Education Improvement Act, Pub. L. No. 108-446, 118 Stat. 2647 (2004).
- Institute of Education Sciences. (2005, March). *Biennial report to Congress*. Retrieved from <http://ies.ed.gov/pdf/biennial-rpt05.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](http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_version1_standards.pdf)
- \*Jitendra, A. K., Hoppes, M. K., & Xin, Y. P. (2000). Enhancing main idea comprehension for students with learning problems: The role of a summarization strategy and self-monitoring instruction. *Journal of Special Education, 34*, 127–139.
- \*Johnson, G., Gersten, R., & Carnine, D. (1987). Effects of instructional design variables on vocabulary acquisition of learning disabilities students: A study of computer-assisted instruction. *Journal of Learning Disabilities, 20*, 206–213.
- Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., & Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A practice guide* (NCEE 2008-4027). Washington, DC: U.S. Department of Education, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences. Retrieved from <http://ies.ed.gov/ncee/wwc>
- \*Kennedy, K. M., & Backman, J. (1993). Effectiveness of the Lindamood auditory in depth program with students with learning disabilities. *Learning Disabilities Research & Practice, 8*, 253–259.
- \*Kim, A., Vaughn, S., Klingner, J. K., Woodruff, A. L., Reutebuch, C., & Kouzekanani, K. (2006). Improving the reading comprehension of middle school students with disabilities through computer-assisted collaborative strategic reading. *Remedial and Special Education, 27*, 235–249. doi:10.1177/07419325060270040401
- \*Kim, J. S., Samson, J. F., Fitzgerald, R., & Hartry, A. (2010). A randomized experiment of a mixed-methods literacy intervention for struggling readers in grades 4–6: Effects on word reading efficiency, reading comprehension and vocabulary, and oral reading fluency. *Reading and Writing, 23*, 1109–1129. doi:10.1007/s11145-009-9198-2
- \*Klingner, J. K., & Vaughn, S. (1996). Reciprocal teaching of reading comprehension strategies for students with learning disabilities who use English as a second language. *Elementary School Journal, 96*, 275–293.
- \*Lang, L., Torgesen, J., Vogel, W., Chanter, C., Lefsky, E., & Petscher, Y. (2009). Exploring the relative effectiveness of reading interventions for high school students. *Journal of Research on Educational Effectiveness, 2*, 149–175. doi:10.1080/19345740802641535
- \*Lovett, M. W., Lacerenza, L., De Palma, M., & Frijters, J. C. (2012). Evaluating the efficacy of remediation for struggling readers in high school. *Journal of Learning Disabilities, 45*, 151–169. doi:10.1177/0022219410371678
- \*Macaruso, P., & Rodman, A. (2009). Benefits of computer-assisted instruction for struggling readers in middle school. *European Journal of Special Needs Education, 24*, 103–113. doi:10.1080/08856250802596774
- \*Manset-Williamson, G., & Nelson, J. M. (2005). Balanced, strategic reading instruction for upper-elementary and middle school students with reading disabilities: A comparative study of two approaches. *Learning Disability Quarterly, 28*, 59–74. doi:10.2307/4126973
- \*Mastropieri, M. A., Scruggs, T. E., Levin, J. R., Gaffney, J., & McLoone, B. (1985). Mnemonic vocabulary instruction for learning disabled students. *Learning Disability Quarterly, 8*, 57–63.
- \*Mastropieri, M. A., Scruggs, T., Mohler, L., Beranek, M., Spencer, V., Boon, R. T., & Talbot, E. (2001). Can middle school students with serious reading difficulties help each other and learn anything? *Journal of Learning Disabilities, 16*, 18–27.
- \*McCallum, R., Krohn, K. R., Skinner, C. H., Hilton-Prillhart, A., Hopkins, M., Waller, S., & Polite, F. (2011). Improving reading comprehension of at-risk high-school students: The ART of Reading program. *Psychology in the Schools, 48*, 78–86. doi:10.1002/pits.20541
- \*McLoone, B. B., Scruggs, T. E., Mastropieri, M. A., & Zucker, S. F. (1986). Memory strategy instruction and training with learning disabled adolescents. *Learning Disabilities Research, 2*, 45–52.
- \*Meyer, B. F., Wijekumar, K. K., & Lin, Y. (2011). Individualizing a web-based structure strategy intervention for fifth graders' comprehension of nonfiction. *Journal of Educational Psychology, 103*, 140–168. doi:10.1037/a0021606
- \*Moore, P. J., & Scevak, J. J. (1995). The effects of strategy training on high school students' learning from science texts. *European Journal of Psychology of Education, 10*, 401–410.

- National Center for Education Statistics. (2011). *The nation's report card: Reading 2011* (NCES 2012-457). Washington, DC: U.S. Department of Education, Institute of Education Sciences.
- \*O'Shea, L. J., Sindelar, P., & O'Shea, D. J. (1987). Effects of repeated readings and attentional cues on the reading fluency and comprehension of learning disabled readers. *Learning Disabilities Research, 2*, 103–109.
- \*Penney, C. G. (2002). Teaching decoding skills to poor readers in high school. *Journal of Literacy Research, 34*, 99–118.
- \*Rasinski, T., Samuels, S., Hiebert, E., Petscher, Y., & Feller, K. (2011). The relationship between a silent reading fluency instructional protocol on students' reading comprehension and achievement in an urban school setting. *Reading Psychology, 32*, 75–97. doi:10.1080/02702710903346873
- Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C. K., . . . Torgesen, J. K. (2007). *Reading interventions for adolescent struggling readers: A meta-analysis with implications for practice*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.
- \*Shippen, M. E., Houchins, D. E., Steventon, C., & Sartor, D. (2005). A comparison of two direct instruction reading programs for urban middle school students. *Remedial and Special Education, 26*, 175–182. doi:10.1177/07419325050260030501
- \*Snider, V. E. (1989). Reading comprehension performance of adolescents with learning disabilities. *Learning Disability Quarterly, 12*, 87–96.
- \*Somers, M., Corrin, W., Sepanik, S., Salinger, T., Levin, J., & Zmach, C. (2010). *The enhanced reading opportunity study final report: The impact of supplemental literacy courses for struggling ninth-grade readers*. Washington, DC: U.S. Department of Education, Institute of Educational Sciences, National Center for Education Evaluation and Regional Assistance.
- \*Spencer, S., & Manis, F. (2010). The effects of a fluency intervention program on fluency and comprehension outcomes of middle-school students with severe reading deficits. *Learning Disabilities Research & Practice, 25*, 76–86. doi:10.1111/j.1540-5826.2010.00305.x
- Swanson, H. L., Hoskyn, M., & Lee, C. (1999). *Interventions for students with learning disabilities: A meta-analysis of treatment outcomes*. New York, NY: Guilford.
- \*Thames, D. G., Reeves, C., Kazelskis, R., York, K., Boling, C., Newell, K., & Wang, Y. (2008). Reading comprehension: Effects of individualized, integrated language arts as a reading approach with struggling readers. *Reading Psychology, 29*, 86–115. doi:10.1080/02702710701853625
- \*Therrien, W. J., Wickstrom, K., & Jones, K. (2006). Effect of a combined repeated reading and question generation intervention on reading achievement. *Learning Disabilities Research & Practice, 21*, 89–97. doi:10.1111/j.1540-5826.2006.00209.x
- \*Torgesen, J., Stancavage, F., Myers, D., Schirm, A., Stuart, E., Vartivarian, S., . . . Haan, C. (2006). *Closing the reading gap: First year findings from a randomized trial of four reading interventions for striving readers: Final report*. Washington, DC: U.S. Department of Education, Institute of Education Sciences.
- \*Vadasy, P. F., & Sanders, E. A. (2008). Benefits of repeated reading intervention for low-achieving fourth- and fifth-grade students. *Remedial and Special Education, 29*, 235–249. doi:10.1177/0741932507312013
- \*Vaughn, S., Cirino, P., Wanzek, J., Wexler, J., Fletcher, J., Denton, C., . . . Francis, D. (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., Klingner, J. K., Swanson, E. A., Boardman, A. G., Roberts, G., Mohammed, S. S., & Stillman-Spisak, S. J. (2011). Efficacy of collaborative strategic reading with middle school students. *American Educational Research Journal, 48*, 938–964. doi:10.3102/0002831211410305
- \*Vaughn, S., Wanzek, J., Wexler, J., Barth, A., Cirino, P., 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. doi:10.1007/s11145-009-9183-9
- \*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.
- \*Veit, D. T., Scruggs, T. E., & Mastropieri, M. A. (1986). Extended mnemonic instruction with learning disabled students. *Journal of Educational Psychology, 78*, 300–308.
- \*Wanzek, J., Vaughn, S., Roberts, G., & Fletcher, J. M. (2011). Efficacy of a reading intervention for middle school students with reading disabilities. *Exceptional Children, 78*, 73–87.
- Wanzek, J., Vaughn, S., Scammacca, N., Metz, K., Murray, C., Roberts, G., & Danielson, L. (2013). Extensive reading interventions for students with reading difficulties after Grade 3. *Review of Educational Research, 83*, 163–195.
- Wanzek, J., Vaughn, S., Wexler, J., Swanson, E., Edmonds, M. E., & Kim, A.-H. (2006). A synthesis of spelling and reading interventions and their effects on the spelling outcomes of students with LD. *Journal of Learning Disabilities, 39*, 528–543. doi:10.1177/00222194060390060501
- \*Wexler, J., Vaughn, S., Roberts, G., & Denton, C. A. (2010). The efficacy of repeated reading and wide reading practice for high school students with severe reading disabilities. *Learning Disabilities Research & Practice, 25*, 2–10. doi:10.1111/j.1540-5826.2009.00296.x
- \*Wilder, A. A., & Williams, J. P. (2001). Students with severe learning disabilities can learn higher order comprehension skills. *Journal of Educational Psychology, 93*, 268–278.
- \*Williams, J. P., Brown, L. G., Silverstein, A. K., & deCani, J. S. (1994). An instructional program in comprehension of narrative themes for adolescents with learning disabilities. *Learning Disabilities Quarterly, 17*, 205–221.
- Willingham, D. T. (2007). Ask the cognitive scientist: The usefulness of brief instruction in reading comprehension strategies. *American Educator, 30*(4), 39–45, 50.
- Willingham, D. T. (2012, April 30). *Collateral damage of excessive reading comprehension strategy instruction* [Blog post]. Retrieved from <http://www.danielwillingham.com/1/post/2012/4/collateral-damage-of-reading-comprehension-strategy-instruction.html>