

**What We Know and Need to Know about Literacy Interventions for Elementary Students  
with Reading Difficulties and Disabilities, including Dyslexia**

**Stephanie Al Otaiba<sup>1</sup>**

**Kristen McMaster<sup>2</sup>**

**Jeanne Wanzek<sup>3</sup>**

**Mai W. Zaru<sup>1</sup>**

**DOI:** 10.1002/rrq.458

1 Southern Methodist University, Department of Teaching and Learning, Dallas, TX

2 University of Minnesota, Department of Educational Psychology, Minneapolis, MN

3 Vanderbilt University, Department of Special Education, Nashville, TN

*The research reported here was supported by the Institute of Education Sciences, U.S.*

*Department of Education, under Award Numbers R324A160132 and R324A170101, and by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R01HD091232-01A1. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Institute for Education Sciences or the National Institutes of Health.*

Ethical approval was not required because this literature review involved retrieval and synthesis of data from existing reviews and meta-analyses of previously published studies.

Accepted online 2021

Correspondence should be addressed to Stephanie Al Otaiba, Southern Methodist University,  
Department of Teaching and Learning, PO Box 750455, Dallas, TX 75025-0455. Email:

[salotaiba@smu.edu](mailto:salotaiba@smu.edu).

### Abstract

The purpose of this paper is to describe what we know and what we still need to learn about literacy intervention for children who experience significant difficulties learning to read. We reviewed 14 meta-analyses and systematic reviews of experimental and quasi-experimental studies published in the last decade that examined the effects of reading and writing interventions in the elementary grades, including research focused on students with reading difficulties and disabilities, including dyslexia. We attended to moderator analyses, when available, to further refine what we know and need to learn about interventions. Findings from these reviews indicate that explicit and systematic intervention focusing on the code *and* meaning dimensions of reading and writing, and delivered one-to-one or in small groups, are likely to improve foundational code-based reading skills, and to a lesser extent, meaning-based skills, across elementary grade levels. Findings, at least in the upper elementary grades, indicate that some intervention features including standardized protocols, multiple components, and longer duration can yield stronger effects. And, integrating reading and writing interventions shows promise. We still need to learn more about specific instructional routines and components that provide more robust effects on students' ability to comprehend and individual differences in response to interventions. We discuss limitations of this review of reviews and suggest directions for future research to optimize implementation, particularly to understand *for whom* and *under what conditions* literacy interventions work best.

Keywords: *literacy intervention, dyslexia, struggling readers, reading difficulty, reading disability*

**What We Know and Need to Know about Literacy Interventions  
for Elementary Students with Reading Difficulties and Disabilities, including Dyslexia**

The literature addressing what we know about literacy interventions to support elementary students who experience significant difficulties learning to read is broad and deep, spanning several decades of experimental and quasi-experimental research (e.g., National Reading Panel, 2000). Despite this research base, it is concerning that only about a third of fourth graders achieve proficient reading levels (National Center for Education Statistics, U.S. Department of Education, 2019). Among students identified with a specific learning disability (SLD) only 33% perform at even a basic level (NCES, 2019). These data reflect only those students with dyslexia who have been also identified with a SLD under the Individuals with Disabilities Education Act, so it does not represent reading levels for all students with dyslexia, given that some may receive support through the Section 504 of the Rehabilitation Act (1973) and some may not have received a school-assigned dyslexia classification.

The extant literature offers valuable insights regarding what constitutes effective literacy intervention for students identified with reading difficulties and disabilities, including dyslexia. Researchers have synthesized much of this work into meta-analyses and systematic reviews that provide guidance for research and practice. In this paper, we aggregate findings from these syntheses into a ‘review of reviews’ to shed further light on what we know and still need to learn about literacy intervention for children who experience significant difficulties learning to read. We synthesize meta-analyses and systematic reviews published in the past decade to reflect a timeframe when the vast majority of states have mandated dyslexia screening and intervention supports (see introduction for this special issue).

## Conceptual Framework

The *Simple View of Reading* (Gough & Tunmer, 1986) describes the necessary components of reading skills, which in turn help identify the types of difficulties that readers who struggle might encounter, and the types of skills teachers and interventionists should emphasize during instruction. Specifically, reading is the product of decoding, or *code-focused skills* and linguistic comprehension, or *meaning-focused skills*. Code-focused skills include phonemic awareness, letter-sound correspondence, phonics (including decoding and encoding, or spelling) and word recognition. Meaning-focused skills include vocabulary, oral language, and listening and reading comprehension skills. This *Simple View of Reading* provides a foundation for the conceptual framework within which we contextualize the need for intensive interventions designed to support children who experience significant difficulties learning to read and those with reading disabilities, including dyslexia.

The *Simple View of Reading* (Gough & Tunmer, 1986) implies four subgroups of readers. Our review does not focus on research syntheses related to the subgroup of *good readers* who have developed both set of skills, nor did we focus on the subgroup of students who experience difficulties only with meaning skills, or *poor comprehenders*. Rather, we focus on research evaluating the effectiveness of interventions for the subgroups of students who experience difficulties with code-focused skills, but have adequate language comprehension, or *dyslexia*, and who experience difficulties with both code- and meaning-focused skills. We included both of these subgroups because researchers have evaluated interventions for students with a wider array of students at risk for or with reading difficulties or disabilities, without specifically naming ‘dyslexia’ as a condition—reflecting, perhaps, a lack of consensus in the field regarding how to operationalize and label significant reading problems (Elliott, 2020; Grigorenko et al., 2020;

Miciak & Fletcher, 2020; Seidenberg et al., 2020; Snowling et al., 2020). Significant reading difficulties stemming from low achievement in reading and spelling, with related impacts on comprehension, are likely to persist without targeted interventions (Catts et al., 2006; Cervetti et al., 2020) and may lead to a persistent pattern reflecting intractable inadequate response to intervention (e.g., Miciak & Fletcher, 2020). Students identified with dyslexia, those with primarily word-level reading difficulties, not only struggle to develop key skills related to cracking the code that allows them to decode and spell words, but also struggle to read with fluency, limiting attentional capacity needed to comprehend text (LaBerge & Samuels, 1974; Samuels et al., 2005). Given that the ultimate goal of reading is to read with understanding (Castles et al., 2018; Ehri, 2005, 2020), word-level difficulties, if not adequately addressed, can lead to more pervasive challenges related to language development and to comprehending and learning from text, which in turn can lead to significant and long-term negative consequences that impact students' academic and lifelong success (Elliott & Grigorenko, 2014; Snow & Biancarosa, 2003).

In addition to describing the difficulties that readers might experience, the *Simple View of Reading* implies that reading instruction and interventions should include both code- and meaning-focused instruction. Schools typically provide such reading intervention within multi-tiered systems of support (MTSS), or Response to Intervention (RTI) in which students who need more targeted support than that offered in core instruction (Tier 1) receive supplementary intervention (Tier 2) (e.g., Gersten et al., 2009; Vaughn, et al., 2003). Supplementary intervention often takes the form of validated standard treatment protocols delivered in small groups by an intervention specialist who monitors student responsiveness using progress-monitoring data. Those students for whom supplementary intervention is insufficient—including

many students at risk or identified with dyslexia or other reading disabilities—may require intervention that is even more intensive (Tier 3).

### ***Effective Reading Interventions***

In their review of the research on multi-tiered interventions in the primary grades, Gersten et al. (2009) included a recommendation for providing explicit systematic instruction (overtly teaching each step of a process, including modeling) for up to three reading skills within the code- and meaning-focused domains (e.g., phonemic awareness, phonics, fluency, vocabulary and comprehension). A strong level of evidence (i.e., evidence-based) signifies that the studies had both high internal validity, indicating strong research designs that were experimental or quasi-experimental, and high external validity, indicating there were enough studies to generalize across participants and settings. Based on their review, Gersten and colleagues recommended these interventions should be implemented in small homogeneous groups at least three times per week for 20-40 min sessions. They indicated that, whether teachers or paraprofessionals provided these interventions, it was important to build skills gradually and provide high levels of support with opportunities for students to practice and receive feedback. However, they noted a low level of evidence to inform more intensive interventions (i.e., Tier 3) for students who had demonstrated minimal progress (“nonresponders”) after a reasonable time in Tier 2 small group intervention. They called for additional research on interventions for these students.

Although research on reading interventions in the upper elementary grades is more limited than research in the early grades, several syntheses of reading intervention research after Grade 3 have been conducted in the past decade. Most of the syntheses have combined upper elementary grades (Grade 4 and 5) with middle and secondary grades. A review of research on

adolescent literacy instruction (Grades 4-12) by Kamil et al. (2008) included a recommendation for implementation of intensive, individualized interventions for struggling readers provided by qualified specialists. Kamil et al. (2008) indicated the research evidence for this recommendation was strong, meaning studies with high internal and external validity provided positive findings for the recommendation. The research-based recommendation included evidence for an explicit instructional focus and greater intensity of instruction for students with greater instructional needs, with intensity largely defined as smaller instructional groups and/or more instructional time.

There is a need to better understand the more recent evidence for reading intervention, particularly with the rapid proliferation of dyslexia legislation and policies over the past decade, the use of multi-tiered systems of support and screening for identification, and recommendations about interventions (e.g., Petscher et al., 2020; Youman & Mather, 2018). Many states have specified that schools conduct dyslexia-specific screening and implement dyslexia-specific intervention on top of existing efforts, which may be challenging for under-resourced schools struggling to provide students with effective instruction and intervention. The National Center for Improving Literacy (National Center on Improving Literacy, 2020) provides updates regarding states' recommendations for intervention policies, which currently include the following terms: "multi-sensory," "evidence-based," "explicit/direct," and/or "MTSS/RTI." Multisensory refers to instruction that incorporates teaching through the senses, or modalities (visual, auditory, tactile, and kinesthetic), to train and reinforce connections between spoken and written language (e.g., Birsch, 2011). Briefly, the origins of multi-sensory approaches are nearly a century old and include arguments that students with dyslexia need not only explicit

instruction, but also need support to connect their visual, auditory, tactile, and kinesthetic modalities (e.g., Birsch, 2011; Orton, 1937).

### ***Learning to Write to Support Learning to Read***

The present review primarily focuses on intensive reading interventions; yet, we believe this work would be incomplete without recognizing the critical role that writing plays within children's overall literacy development, and reading outcomes specifically. Theoretical and empirical evidence points to a strong relation between reading and writing (e.g., Berninger et al., 2002; Fitzgerald & Shannahan, 2000), and researchers have estimated that about three-fourths of students who experience significant writing difficulties have comorbid reading difficulties. Further, Graham et al. (2020) conducted a meta-analysis to determine whether children identified with reading difficulties and disabilities also experience writing difficulties, and found substantial evidence that, indeed, they do. It is likely that many students who experience reading difficulties or disabilities—including those with dyslexia—could benefit from writing instruction, and that effective writing instruction in turn could improve their reading development.

Researchers (e.g., Graham, 2020; Shanahan, 2016) have explained the theoretical relation between reading and writing from a “shared knowledge” perspective—positing that reading and writing both draw on the same knowledge and cognitive systems. These knowledge and cognitive systems include general background knowledge needed for both generating and comprehending text, meta-knowledge about written language systems (such as genre and text structure), procedural knowledge about how to interact with text, and pragmatic knowledge about text attributes (such as alphabetic knowledge needed to both decode and encode text; knowledge about words, syntax, and so on). For students with word-level reading difficulties,

including those with dyslexia, alphabetic knowledge needed to encode text (i.e., spelling skills) is particularly important to both reading and writing development. In fact, Ehri (2000) pronounced spelling and word-reading skills to be “two sides of a coin” (p. 19) given that both rely on letter-sound knowledge (with word reading requiring decoding of words based on letter-sound knowledge, and spelling requiring *encoding* of words based on letter-sound knowledge). Empirical evidence supports this relation: researchers have reported a strong correlation between reading and writing skills for students with learning disabilities (e.g., Berninger et al., 2002), and have provided evidence that spelling can enhance phonemic awareness, alphabetic awareness, and sight word recognition (Adams, 1990; Conrad, 2008; Ehri, 1987, 2005; Moats, 2005 ; Treiman, 1993).

In light of this “shared knowledge” perspective and evidence that students who experience significant reading difficulties also often struggle with writing, we include meta-analyses that examine effects of writing instruction (including spelling instruction) on reading outcomes in our review. Our aim in doing so is to shed light on the role that writing instruction can have on reading outcomes, as integrating reading and writing instruction might be an important way to increase educators’ capacity to meet the needs of children with reading difficulties and disabilities. In other words, if evidence indicates that strengthening students’ skills in writing also strengthens reading skills, then researchers should consider exploring writing instruction as a component of intensive literacy intervention.

### **Purpose and Organization**

The purpose of this paper is to describe what we know and what we still need to learn about literacy intervention for elementary-aged students who experience significant difficulties learning to read. Given legislation enacted over the past decade in states across the U.S.

addressing the need to ensure effective instructional programming in schools for children identified with dyslexia (National Center on Improving Literacy, 2020; Petscher et al., 2020), we synthesize meta-analyses and systematic reviews published in the last 10 years. To achieve this aim, we review meta-analyses and systematic reviews published in the last decade that focus on reading interventions for students with reading difficulties or disabilities, including dyslexia in the elementary grades. We also explore existing reviews on the effects of writing interventions on reading outcomes for children with reading difficulties, given important theoretical and empirical relations between reading and writing. In doing so, we attend to moderator analyses, when available, in order to describe individual differences in response, and identify specific instructional features, including intensity of intervention, that emerge as particularly relevant. Ethical approval was not required because this literature review involved retrieval and synthesis of data from existing reviews and meta-analyses of previously published studies.

Specific research questions that guided our review are as follows: (1) What are the effects of early reading interventions conducted in Grades K-3? (2) What are the effects of reading interventions conducted in the upper elementary grades? (3) What are the effects of writing instruction on reading skills across the elementary grades? We address each question in the following sections, and then discuss what we know and still need to learn about effective literacy instruction for children with reading difficulties and disabilities, including dyslexia.

### **Review of Reviews**

To address our research questions, we searched the ERIC and PsychINFO databases for peer reviewed journal articles published from 2010-2020 using the following search terms: (a) *meta-analyses, literature review, meta-analysis*, AND (b) *reading risk, reading difficulty, learning disab\*, non-responder*, AND (c) *intervention, reading intervention, systematic*

*instruction, structured literacy, multisensory, or Orton-Gillingham*. In the second step, we added the search term *writing interventions* and then in the third step, we added the term *dyslexia*. We included all meta-analyses and literature syntheses that included students in the elementary grades; we did not exclude studies that also spanned beyond fifth grade, in order to capture effects for upper elementary students. We also identified two relevant meta-analyses that were in review. We focused our descriptions of the effects for participants in the elementary grades. Overall, 14 syntheses were identified in this search.

### **Syntheses and Meta-analyses Describing the Effectiveness of Early Reading Interventions Conducted in Grades K-3**

We identified one best evidence synthesis, four meta-analyses, and two narrative syntheses conducted in the past decade that report the effectiveness of intensive early reading interventions provided in kindergarten through third grade which validate and extend Gersten et al.'s (2009) recommendations for implementing multi-tier interventions for struggling readers. We summarize these studies in chronological order. Table 1 provides an overview of the key aspects of each research review.

#### ***Effects of Interventions in Grades K-3***

**Slavin et al. (2011).** Slavin and colleagues used a best-evidence synthesis approach to examine the research on effective reading programs for struggling readers in Grades K-5. Their inclusion criteria for studies required that participants included students with reading disability, students performing at or below the 33<sup>rd</sup> percentile in reading performance relative to classroom peers, or students receiving intensive reading services. They included rigorous studies that used randomization to condition (experimental or quasi-experimental studies) or appropriate matching; they also included only studies that provided pre- and post-treatment data on

standardized or distal experimenter-created measures. At the minimum, studies must have had a 12-week duration with a minimum of 15 students and two tutors per group. The broad scope of their synthesis included some types of instructional programs that are not directly relevant to our present review, which we do not discuss (i.e., classroom instructional programs for all students or computer-assisted instruction).

We focused on Slavin et al.'s (2011) findings describing the effects of three types of intensive interventions: (1) one-to-one tutoring by certified teachers, (2) one-to-one tutoring by paraprofessionals and volunteers, and (3) small-group tutoring, which represented 58 studies. Of these, 20 studies examined one-to-one tutoring by certified teachers, which Slavin and colleagues termed the "gold standard" for intervention delivery. The vast majority of these 20 studies were conducted in the primary grades; only one study included students in upper elementary in their sample (including grades 2-6). For these 20 studies, the overall weighted mean effect size (ES) for reading outcomes was 0.39. Although they did not conduct a moderator analysis, Slavin and colleagues noted that the weighted mean ES for the subset of 8 studies involving a less explicit and systematic approach, *Reading Recovery*, (Pinnell, et al., 1988) was relatively smaller (0.23). By contrast, the remaining 12 studies, in which intervention included an explicit and systematic approach to phonics, had a larger weighted mean ES of 0.56.

The second type of one-to-one tutoring studies they reviewed were provided by paraprofessionals and volunteers. A majority of these 18 studies took place in K-1 classrooms. Only two spanned primary and upper elementary, and only one was solely conducted in upper elementary. The overall mean weighted ES for these 18 studies for reading outcomes was 0.24. Slavin et al. compared the effect sizes in tutoring studies delivered by paraprofessionals to volunteers, which were 0.38 and 0.16, respectively, suggesting weaker effects of interventions

provided by volunteers. They also noted that the interventions were standardized programs that emphasized code-focused instruction.

The third type of studies Slavin and colleagues (2011) reviewed were small group tutorials, which included 20 studies examining 18 different programs. Similar to the tutoring studies, the interventions were standardized tutorial programs that included training for the interventionists, structured materials, and an emphasis on phonics. A majority of the studies focused on primary grade students, with eight studies that spanned primary and upper elementary grades. The overall mean weighted ES for reading outcomes was 0.31.

**Wanzek et al., (2015).** To further examine effects of reading interventions provided to struggling readers in kindergarten through third grade, Wanzek et al. conducted four meta-analyses that included 72 studies. Across studies, participants were identified as at risk or with reading difficulties; four included students identified as having a reading disability and one included students with speech language disabilities. The majority of studies ( $n = 48$ ) focused on students in kindergarten and first grade, with the remainder including students in multiple grade levels or students in second and third grade.

The majority (63 of the 72 studies) examined the effect of code-focused reading skill interventions, which included phonological awareness, word reading, and fluency. These interventions ranged in duration from 15-99 sessions. Wanzek et al. (2015) reported an overall average ES of 0.54 on standardized measures of code-focused reading skills, which is similar to the ES of 0.56 reported by Slavin et al. (2011) for the subset of teacher-led explicit and systematic tutoring. Within the 33 studies that included non-standardized measures of code-focused reading skills, the mean ES was 0.62. Thus, this large meta-analysis confirmed a moderate and positive effect of interventions on code-focused reading skills. The overall average

ES of interventions on meaning-focused skills (language and comprehension) were smaller for standardized measures (0.36) and larger for non-standardized measures (1.02).

There were enough studies to allow authors to conduct analyses to determine whether there were statistically significant differences in effects based on key moderator variables, but only for foundational code-focused reading skills. None of the moderator analyses were significant (see Table 1). Thus, an important implication of this meta-analysis is that interventions had similar positive and moderate effects regardless of intervention type, grade level, implementer (general education, special education teachers, researchers, or paraprofessionals), group size, or number of hours of intervention.

**Austin et al. (2017).** Among the studies we reviewed, Austin and colleagues were the only research team to focus on the effect of Tier 3 interventions provided to elementary students who had previously demonstrated inadequate response to less-intensive (Tier 2) intervention. They located a total of 12 experimental and quasi-experimental studies conducted in Grades K-3. We have not incorporated their findings into Table 1 because this was a narrative synthesis that did not provide an overall effect size. One challenge researchers faced in comparing effects of intensive interventions across these 12 studies was the variability in how study authors operationally defined “inadequate response” to intervention. Most researchers used a criterion for post-test performance on a standardized reading measure, but they interpreted this criterion differently. Two studies evaluated response relative to scores for a local peer group, three used *z*-scores summed to determine a rank order, and two used a percentile score cut-point on one or more measures. Three studies used benchmarks or cut-points for oral word reading fluency on criterion-referenced curriculum based measures, and one relied on classroom teacher judgment.

Another challenge to comparing findings and effects in the Austin et al. (2017) synthesis was that researchers employed different types of counterfactual or comparison conditions. In the majority of studies, the students in the comparison condition were higher performing students, including students who had responded adequately to the Tier 2 intervention, students who were at grade level and were never eligible for intervention, or a combination of the two. Among these studies, students who received more intensive interventions did not sufficiently accelerate their growth to catch up to students in the higher performing control groups. Three studies recruited students who had demonstrated inadequate response to previous intervention and then randomly assigned them to more intensive intervention(s) or to a business as usual control condition. Findings from these three studies documented significantly stronger outcomes for students who received an intensive explicit and systematic Tier 3 relative to those in the control condition. We review these studies in more detail here given their particular relevance to the present paper in describing what we know and need to know about intensive interventions.

The first study (Scanlon et al., 2005) was conducted across first grade by certified teachers who provided daily sessions of intervention to individual students who had not responded adequately in kindergarten. They provided two contrasting interventions relative to a control condition. Both interventions included phonological skill, word instruction, repeated reading for fluency, and writing; but they differed in time allocation to phonological skill vs text reading. Students in both treatment conditions demonstrated significantly greater growth than controls on the majority of reading measures, with ES for basic word reading of 0.58 to 0.64, and for comprehension of 0.07 to 0.42.

The second study, conducted by Denton et al. (2006), was briefer in duration, but students received a “double dose” of daily intervention. Teachers provided systematic and

explicit phonics intervention to pairs of students over 8 weeks for two daily 50-minute sessions. The students' grades ranged from first through third. The students in the intervention condition demonstrated significantly greater growth than those in the control condition, with ES ranging from 0.39 to 1.77 for word reading; and from 0.37 to 0.63 for comprehension.

In the third study, Denton et al. (2013) provided intervention for four to five 45 min weekly sessions in small groups of 2-3 for between 24-26 weeks. The treatment involved an adaptive intervention that provided phonics, fluency, and comprehension instruction, and the implementors used data-based decision making to adapt lessons for student need. Their students were either in second grade or were repeating first grade. The students in the adaptive intervention condition demonstrated significantly greater growth than those in the control condition, with ES ranging from 0.39 to 0.56 for word reading and from 0.34 to 0.35 for comprehension.

The results of these three studies support the efficacy of more intensive intervention for students who had not demonstrated response to prior intervention by providing more sessions per week, decreasing group size, increasing session length or duration; and in the last study, by adapting lessons based on student progress data.

**Al Otaiba et al. (2018).** Al Otaiba and colleagues wrote a practitioner-friendly synthesis of the literature describing elementary-grade intervention approaches for students with intensive reading and writing needs, including students with dyslexia. They extended the one previously published synthesis (Ritchey & Goeke, 2006) describing effects of multi-sensory Orton-Gillingham (OG) approaches relative to a control group. There were some positive effects of OG interventions on standardized measures for word reading, with a mean ES of 0.42, but in the six studies that provided effect sizes, these effects varied from large and positive (1.56) to large and

negative (-0.91). In at least two studies, findings indicated that the OG approaches were less effective than the comparison group and another study reported no significant differences.

Al Otaiba et al. (2018) located only one experimental study that directly compared the efficacy of an explicit and systematic intervention to one that also incorporated a multi-sensory approach (Schlesinger & Grey, 2017). The researchers used a multiple probe, multiple baseline single case design. Schlesinger and Grey (2017) They reported that both approaches showed important treatment effects over baseline; however, there were no significant effects favoring multi-sensory methods over the explicit and systematic approach on letter naming, letter sound production, word reading, or word spelling outcomes.

Al Otaiba et al. (2018) also examined findings from three reports from the Institute for Education Science What Works Clearinghouse (WWC) to learn whether additional, more recent technical reports or studies provided stronger evidence about widely used multi-sensory interventions. The WWC concluded that there were not enough strong studies to determine whether OG approaches were evidence based (WWC, 2010<sup>a</sup>). They did report findings from the few existing randomized control trials for other widely-used programs that incorporate multi-sensory approaches including, for example *Lindamood Phoneme Sequencing (LiPS; Lindamood, 1998)* and *Wilson Reading (1988)* (WWC, 2010<sup>b</sup>; WWC, 2010<sup>c</sup>). The report on *LiPS* indicated mixed findings: one study (Gunn, 1996) reported mean negative ES on word reading (-0.33), whereas the other (Torgesen et al., 2010) reported significant and positive mean ES on standardized measures of word reading and phonological awareness (0.63) and comprehension (0.54). The report on *Wilson* revealed a significant and positive effect (mean ES of 0.33 for word reading) (Torgesen et al., 2006). Al Otaiba et al. (2018) concluded that more studies with stronger internal and external validity were needed to develop an evidence base for the efficacy

of dyslexia-specific approaches, and also to directly compare these approaches to other explicit and systematic reading and writing interventions that met the WWC criteria for strong evidence for struggling readers. We have not incorporated findings from this narrative synthesis into Table 1 because it did not provide an overall ES.

**Wanzek et al. (2018).** Wanzek and colleagues conducted a meta-analysis of 25 studies that examined effects of interventions that provided intervention for longer durations - all interventions lasted a minimum of 100 sessions. The authors selected studies that included students in grades K-3 identified with a learning disability or as at risk for reading difficulty, or struggling readers. Three studies included students with disabilities. Wanzek et al. (2018) reported the weighted mean ES for reading was 0.39, which was slightly lower after adjustments were made for publication bias (0.28). Their findings extend findings of prior studies in this review and further confirm the moderate positive effect of intervention on reading skills.

To understand whether pretest scores on standardized reading tests predicted effect sizes, Wanzek et al. (2018) conducted a meta-regression, which revealed no significant prediction. In other words, students with initially weaker reading scores responded similarly to other participants. However, they cautioned that only 12 of the 25 studies provided pretest data. They suggested that future research should report pretest scores on standardized measures to better inform how to match student needs to various types of intensive interventions. The researchers also conducted moderator analyses, but as in the prior meta-analyses, found no significant heterogeneity in effect sizes based on intervention type, grade level, implementer, group size, or hours of intervention. They cautioned that their moderator analyses might have had relatively low power; and so, given their exploratory nature, they reported effect sizes. They found similar ES for researcher and teacher implemented interventions (0.52 and 0.50, respectively). Similar to

Slavin et al. (2011), they found one-to-one tutoring tended to have higher ES (0.59) than small group interventions (0.33) and ES were higher for interventions lasting more than 63 hours (0.45) than those lasting 63 hours or less (0.33).

**Gersten et al. (2020).** The most recent meta-analysis of the effects of reading interventions provided to struggling readers focusing only on the primary grades was conducted by Gersten et al. (2020). Their study differed from prior studies in that they selected only those studies that met the standards for methodological rigor based on WWC standards as meeting criteria for “moderate” to “strong” levels of evidence. They did so because the Every Student Succeeds Act (ESSA, 2015) called for an emphasis on interventions meeting this criteria as evidence-based. Gersten and colleagues reviewed 33 studies conducted in Grades 1-3, with the majority ( $n = 22$ ) conducted in Grade 1, and with 11 studies conducted in Grades 2 or 3. Given this grade range, they chose to exclude studies that provided only phonological awareness training without providing word-reading instruction. Unlike Austin et al., 2017, they specifically excluded any studies that provided Tier 3 intensive interventions, which Gersten and colleagues defined as interventions provided to students who had not responded to Tier 2 small group intervention.

Gersten et al. categorized participants into two levels of risk based on performance on a standardized norm-referenced screener: (1) at risk, scoring in the 25<sup>th</sup> percentile or lower or (2) minimal risk, scoring below the 40<sup>th</sup> percentile. The interventions had to last for a minimum of eight hours (which they categorized into three dosage levels: (1) low, less than one and a half hours per week, (2) medium, one and a half to two hours per week, or (3) high two or more hours per week). Interventionists included certified teachers, paraprofessionals, and research staff. They reported an overall average ES for reading interventions on reading outcomes of 0.39. They

also reported the average ES for word reading skills, comprehension, and fluency of 0.41, 0.32, and 0.31, respectively. These ES are consistent with those reported by Slavin et al. (2011), and slightly smaller than those reported by Wanzek et al. (2015).

Consistent with moderator analyses conducted by Wanzek et al. (2015) and meta-regression by Wanzek et al. (2018), the results from Gersten et al.'s (2020) moderator analyses that revealed similar effects regardless of grade level, group size, and the level of training for implementers. In addition, student risk, type of comparison group, and study design did not impact effects, nor did intervention characteristics (degree of scripting, whether teachers could modify instruction, group size, or whether the focus was more in-depth and explicit). However, they cautioned that 12 of the 15 scripted programs were provided by paraprofessionals. They did find that intervention type was important: those code-focused interventions that included spelling or writing had stronger ES than those with phonological awareness. They also noted moderation related to the outcome measurement domain, with ES being significantly larger in the word reading, or code-focused, domain compared to either the comprehension or fluency within the meaning-focused domain.

**Stevens et al. (2021).** The purpose for Stevens and colleagues' meta-analysis was to summarize the research examining the effects of multi-sensory Orton-Gillingham (OG) reading interventions in light of the rise in dyslexia-specific legislation, with many states requiring or recommending this type of reading approach. They located 24 studies conducted in Grades K-12 published through 2019 that included OG-branded and unbranded approaches and that had study designs with some type of control condition, including experimental, quasi-experimental, or single case designs. This meta-analysis differed from others we reviewed because it included not only peer-reviewed journal articles, but also dissertations. Of the 24 studies, 8 could not be

included in the meta-analysis, because they did not have a sufficient sample size or had insufficient information to calculate effect sizes. Of the remaining 16 studies, the majority included students in elementary school and only three were conducted with participants outside of the age range for our review (i.e., in middle and high school). Stevens et al found no statistically significant differences in favor of the OG-approaches, although the mean effect sizes for foundational reading skills was 0.22 and for vocabulary and comprehension outcomes was 0.14. The authors cautioned that the methodological quality of the studies they reviewed was relatively weak (e.g., inconsistent descriptions of business as usual conditions, limited fidelity data, and small sample sizes), and that future research with stronger designs is needed to determine the effectiveness of OG relative to other interventions or typical school practices.

### ***Summary of K-3 Findings***

To summarize, findings from these seven syntheses indicated generally positive moderate effects for interventions on foundational skills, with smaller effects for comprehension, although no significant differences were found for OG interventions. A majority of studies were preventative interventions conducted in kindergarten and first grades, with students at-risk for reading disabilities. None of the moderator or meta-regression analyses revealed differential effects based on initial skills or risk or related to type of interventionist. However, there is a need for more research to test the efficacy of intensive interventions for second and third graders, as well as with students with reading disabilities or identified with dyslexia, and students who have not previously responded adequately to generally effective small group interventions. In addition, the findings highlight the paucity of research that has compared the effect of standardized interventions to individualized interventions guided by data.

## **Syntheses and Meta-analyses Describing the Effectiveness of Reading Interventions in the Upper Elementary Grades**

We located four meta-analyses to date that have specifically focused on the intervention research for older students with reading difficulties or disabilities and have analyzed the research base on these supplemental, intensive interventions mentioned by Kamil et al. (2008). We present these in chronological order and identify what we know about reading interventions and their intensity for upper elementary students from each of the analyses. All of these syntheses have examined students with and at-risk for reading-based learning disabilities, but did not limit their target population to only poor decoders. When the authors provided information on decoding or word reading, we have included that in our descriptions. The research does suggest that the majority of students with reading difficulties have significant difficulties in decoding or fluency even into the middle school grades (Cirino et al., 2012).

**Flynn et al. (2012).** Flynn et al. conducted a selective meta-analysis of only experimental research. Ten studies were included, each examining upper elementary or middle school students with reading disabilities, defined as scores lower than the 25<sup>th</sup> percentile on norm-referenced reading measures. The samples as a whole demonstrated pretest scores that were approximately 2 *SDs* below average on word identification measures and approximately 1.5 *SDs* below average on decoding and fluency measures. Seven of these studies included students in the upper elementary grades, but none were specific to only the upper elementary grades. Thus, the findings represent students with reading disabilities in the upper elementary grades but are not specific to only these students. The authors reported an overall ES of 0.37 on norm-referenced reading measures across studies. The largest ES was noted for norm-referenced reading comprehension measures (mean ES = 0.73), and the smallest effects were noted on norm-

referenced reading fluency measures (mean ES = -0.29). These effect sizes represented interventions averaging 41 sessions at 56 min per session over an average of 10 weeks.

However, an examination of moderators indicated no differences among studies that could be attributed to the type of reading measure (reading comprehension, reading fluency, etc.), focus of instruction (comprehension, phonics, phonics and comprehension, etc.), or length of instruction (number of sessions, length of sessions, or weeks of intervention). Thus, even for these samples of students with very poor decoding and word reading, effects of interventions focused on comprehension instruction did not differ from those of interventions focused on phonics or both phonics and comprehension. The authors also examined 18 instructional variables of the interventions. They noted that, of four studies with moderate ESs on reading outcomes, three shared six common instructional variables (controlling task difficulty, explicit practice, teaching new content, sequencing, skill modeling, and task reduction). All four of these studies also implemented interventions that were 46-60 min long. Ultimately, Flynn and colleagues noted moderate, positive effects on student reading outcomes for reading interventions focused on students with reading disabilities, and variation in effects could not be attributed to organizational (length of instruction) or instructional differences.

**Wanzek et al. (2013).** Wanzek et al. further examined the Grade 4-12 research specific to more extensive interventions—those implemented for 75 or more sessions—for students with reading difficulties. The authors analyzed 19 studies (experimental, quasi-experimental, single group, or single case), all examining students with identified learning disabilities and/or reading difficulties, based on assessment of reading achievement as below grade level. Only 10 studies provided the necessary data to be included in the meta-analysis. Notably, 10 of the 19 studies were published in the last 3 years of the search, suggesting an increase in research of more

extensive interventions for students with reading difficulties and disabilities. Nine of the studies included students in the upper elementary grades, with only two studies restricted to only students in the upper elementary grades. Thus, like Flynn et al. (2012), students in the upper elementary grades are represented in the research from this meta-analysis, but the findings are not specific to those grade levels.

Wanzek et al. (2013) reported a mean ES of 0.10 of extensive interventions on reading comprehension outcomes. This was the only outcome with significant variance to examine moderators. However, effects were similar regardless of instructional group size, number of hours of intervention, or grade level of intervention. Reading fluency (ES = 0.16), word reading (ES = 0.15), word reading fluency (ES = 0.16), and spelling (ES = 0.15) outcomes were also analyzed, all with small, positive effects and no significant variation in outcomes noted across studies. Most studies included in this analysis implemented multi-component interventions with three or more reading components addressed in the intervention. The effect sizes of the extensive interventions (75 or more sessions) were lower overall than the effect sizes reported in Flynn et al. (2012), where interventions averaged 41 sessions.

**Scammacca et al. (2013).** Scammacca et al. conducted an additional meta-analysis examining interventions for struggling readers, defined as low-achieving in reading or having unidentified reading difficulties, dyslexia, and/or reading/learning disabilities, in Grades 4-12. Their analysis included 12 studies (experimental and quasi-experimental) examining students in the upper elementary grades, but only three studies focused solely on students in the upper elementary grades. In addition, none of the three upper elementary studies reported the disability status of the participants.

The mean effect across all studies (Grades 4-12) on reading outcomes was 0.49 with a smaller effect of 0.13 when only standardized reading outcomes were included (Scammacca et al., 2013). However, the authors reported a much larger overall ES of 0.96 for studies from 1980-2004; whereas studies from 2005-2011 had a mean ES of 0.23, and the confidence intervals for the two sets of studies (1980-2004 and 2005-2011) did not overlap. Studies published in the later time frame were more likely to have large numbers of participants, which can increase the precision of the study. This set of studies was also more likely to provide at least 16 hours of intervention before measuring the outcomes of the study. Looking at the more recent set of studies (2005-2011), comprehension interventions had a significantly larger ES than fluency or multicomponent interventions. There was no moderation of reading outcomes for the 2005-2011 set of studies based on students' learning disability status, hours of intervention, whether a teacher or a researcher implemented the intervention, or the grade level of intervention.

All three syntheses reviewed thus far examined a wide range of grade levels (Grades 4-12), though two of them noted no moderation of reading outcomes by grade level groupings suggesting that findings for Grades 4-5 may not be different than older grades. However, notably few studies included students specifically in the upper elementary grade.

**Donegan and Wanzek, 2021.** One recent synthesis examined the reading intervention research (experimental and quasi-experimental) for students with reading difficulties specifically at the upper elementary level. This study examined the reading intervention research for struggling readers, defined as students reading below grade level, at-risk for reading disabilities or difficulties, or identified with reading disabilities. The average standard score for participants in the studies from 1988-2007 on foundational reading skills (word recognition, fluency) was 81, demonstrating samples of students with significant difficulties in these word recognition areas.

The average standard score on foundational reading skills for participants in more recent studies (2008-2019) was 90, indicating less impaired readers than in earlier studies, though still performing around the 26<sup>th</sup> percentile in word recognition areas.

The majority of the interventions synthesized by Donegan and Wanzek (2021) addressed either foundational skills only or reading comprehension only during instruction. However, 20 of the 47 interventions were multicomponent interventions addressing both areas of instruction. All but 4 studies provided the interventions in small groups of 1-7 students, with 20 of those interventions provided in the smallest groups of 1-2 students.

The mean effect of these reading interventions on foundational outcomes was 0.22; however, only the ES (0.83) of unstandardized outcomes was significant (Donegan & Wanzek, 2021). The mean effect for standardized foundational outcomes was 0.09. The estimated mean ES of the interventions on comprehension outcomes was 0.21, but this effect was also larger for the unstandardized measures (0.44) than the standardized measures (0.13). Intervention type did moderate both foundational and comprehension outcomes, with multicomponent interventions demonstrating higher effects. Group size also moderated both types of outcomes; small groups of 4-7 students had higher effects than larger groups or smaller groups of 1-2 students for foundational outcomes; yet, groups of 1-2 students had higher effects on comprehension outcomes. Finally, interventions that were more standardized, which were not adjusted based on student progress, had higher effects than the interventions with some portion of the intervention individualized for both foundational and comprehension outcomes. No moderation of foundational outcomes was noted related to the duration of the intervention, but interventions of a longer duration (30 or more hours) had significantly higher effects on comprehension outcomes. Group size remained a significant moderator of student comprehension outcomes even

when all moderators were placed in the models, with instructional group sizes of 1-2 students having the highest effects.

### ***Summary of Effects in Upper Elementary***

The meta-analysis by Donegan and Wanzek (2021) provided the most information on reading interventions for students with reading difficulties or disabilities at the upper elementary levels. It summarized twenty-five studies with a focus on upper elementary students, 17 of which examined only students in 4<sup>th</sup> or 5<sup>th</sup> grade and the additional 8 studies had a majority of participants from these grade levels. The effects for upper elementary students on standardized outcome measures were similar to what Scammacca et al. (2013) and Wanzek et al. (2013) reported, which is not surprising given both of those studies found no significant moderation of grade level on these outcomes.

### **Meta-Analyses Describing the Effects of Writing Instruction on Reading Outcomes**

Thus far, we have summarized research on the effects of reading interventions designed for students who experience significant reading difficulties, demonstrating the importance of explicit and systematic instruction to improve their reading outcomes. As we noted within our conceptual framework, another instructional domain that is critical to consider for improving outcomes of this group is the area of *writing*, given shared knowledge and cognitive systems that support both processes (Shanahan, 2016).

Whereas fewer researchers have examined effects of writing intervention on reading-related outcomes than have examined effects of reading intervention, particularly for students with reading-related difficulties, an evidence base does exist. We identified three meta-analyses published in the last decade that summarize this evidence base and report effect sizes. Unlike the previous meta-analyses summarized in this review, the writing-related meta-analyses were not

restricted to students with specific reading difficulties or disabilities, but we include them here due to the limited research in this area and the important implications for further research conducted to improve outcomes for struggling readers. Also, similar to the syntheses on reading interventions for upper-elementary students, these meta-analyses of writing also included students in middle and high schools. When possible, we disaggregated findings specific to elementary-level students with literacy-related difficulties to shed further light on the importance of writing instruction on reading outcomes for this group.

### *Effects of Writing Instruction and Intervention on Students' Reading Outcomes*

**Graham and Hebert (2011).** Graham and Hebert conducted a meta-analysis to examine the impact of writing instruction on students' reading outcomes. They reviewed experimental and quasi-experimental studies focusing on students in Grades 1-12 that included (a) a treatment group in which students received instruction in writing and (b) at least one reading measure. Relevant studies to the current review ( $n = 21$ ) fell into three categories: (1) Studies that assessed effects of process writing, text structure, and sentence- or paragraph level skills instruction on reading comprehension for students in grades K-12, with ES of 0.22 ( $k = 12$ ) on norm-referenced measures and 0.27 ( $k = 5$ ) on researcher-designed measures; (2) studies that assessed effects of sentence or spelling instruction on reading fluency in Grades 1-7 ( $n = 5$ ) and yielded an average weighted ES of 0.66; and (3) studies that assessed effects of sentence or spelling instruction on word-reading skills in Grades 1-5 ( $n = 6$ ) and yielded an average weighted ES of 0.62 on word-reading outcomes.

Five of the 21 relevant studies focused on students identified as experiencing difficulties with reading and/or writing (three with poor readers, one with poor readers and writers, one with weak spellers). The three with poor readers examined reading comprehension outcomes, yielding

ES of 0.17 to .30. The other two yielded effect ES of 0.39 to 0.51 on word reading outcomes and 0.32 on fluency outcomes. These findings provide promising evidence that struggling readers and writers are likely to experience reading-related benefits as a result of writing instruction, with somewhat weaker effects than those found in studies that included students across the full range of literacy performance. Given the small number of studies that included students with reading and/or writing difficulties, further research is needed to confirm this conclusion, and to shed further light on effects for elementary-level students with significant reading difficulties.

**Graham and Santangelo (2014).** In another meta-analysis, Graham and Santangelo focused specifically on the role of spelling instruction in improving reading outcomes—specifically, by examining experimental and quasi-experimental studies focusing on the effects of formal spelling instruction on spelling and reading outcomes for students in kindergarten through Grade 12. From a “shared knowledge” perspective (e.g., Graham, 2020; Shanahan, 2016), spelling skills are important not only to overall writing development, but also to reading—in fact, Ehri (2000) pronounced spelling and word-reading skills to be “two sides of a coin” given that both rely on letter-sound knowledge (with word reading requiring decoding of words based on letter-sound knowledge, and spelling requiring *encoding* of words based on letter-sound knowledge). Thus, it seems logical to predict that spelling instruction would positively affect word-reading outcomes.

In their review, Graham and Santangelo (2014) examined whether formal spelling instruction enhanced students’ (a) phonological awareness and (b) reading skills. To be classified as formal spelling instruction, more than half of instructional time needed to be devoted to activities in which students had to recode sounds into letters, and involved written practice while learning individual words, typically with feedback provided on spelling accuracy. Like Graham

and Hebert's (2011) meta-analysis, this study was not restricted to students with specific reading difficulties or disabilities, though they did indicate whether studies included students with significant literacy difficulties.

All studies that examined whether explicit spelling instruction improved students' phonological awareness ( $n = 7$ ) were conducted with students in kindergarten or Grade 1; only one focused on students identified as experiencing significant literacy difficulties. Spelling instruction involved either word study or multi-component spelling programs, and researchers compared these programs to phonological, reading, or unrelated instruction. The average weighted effect size across studies was 0.51, indicating that formal spelling instruction had a positive effect on students' phonological awareness.

Studies that examined whether explicit spelling instruction improved students' reading performance ( $n = 20$ ) were conducted with students in kindergarten through Grade 12. Again, spelling instruction involved either word study or multi-component spelling programs; and researchers compared these programs to phonological, reading, whole language, or unrelated instruction. The average weighted effect size across studies was 0.44. Overall (across studies focusing on students representing the full range of literacy skills), effects of explicit instruction were 0.40 on word-reading, 0.36 on fluency, and 0.66 on reading comprehension outcomes. Moderator analyses indicated that quantity of instruction was important, with more spelling instruction significantly associated with larger reading effects. Type of student, grade, quality of study, and type of measure were not significant moderators of effects. Studies that focused solely on students with significant literacy difficulties ( $n = 5$ ; conducted in Grades K-6) reported effect sizes ranging from -0.04 to 0.61 on word-reading outcomes (most ESs were above 0.30), 0.32 on reading fluency (only one study measured fluency), and 0.21 to 0.35 on general measures of

reading. Thus, findings indicate that formal spelling instruction indeed can be beneficial for improving reading outcomes for students who struggle.

**Graham et al. (2018).** Finally, Graham et al. conducted a meta-analysis to examine the impact of literacy instructional programs that balanced reading and writing (with no more than 60% of total instructional time devoted to one or the other). They included experimental and quasi-experimental studies with participants in preschool through Grade 12. Moderator analyses included the following: degree of integration of reading and writing, proportion of instruction devoted to reading vs. writing, whether professional development was provided, whether outcomes were measured using researcher-made vs. norm-referenced measures, sample size, study design, elementary vs. secondary focus, published vs. unpublished studies, typically-developing vs. students experiencing difficulty with literacy, number of instructional sessions, researcher vs. teacher-delivered instruction, and study quality.

Of the studies that met inclusion criteria ( $n = 47$ ), 38 assessed reading performance of students in preschool through high school. Across these studies, average weighted effect sizes were 0.39 for balanced reading and writing instruction on reading comprehension measures, 0.53 on decoding measures, and 0.35 on reading vocabulary measures. Moderator analyses revealed four variables that predicted variability in ES when other study characteristics were controlled. These variables included (a) whether participants were elementary (ES = 0.25) vs. secondary (ES = 0.51) students, (b) whether instruction was equally balanced (ES = 0.67) vs. heavier on reading (ES = 0.33), (c) whether researchers used norm referenced (ES = 0.28) vs. researcher-made (ES = 0.91) assessments, and (d) whether studies were published (ES = 0.36) vs. unpublished (ES = 0.24). Teachers delivered instruction in the majority of studies, and most studies included researcher-provided professional development.

Of the 38 studies with reading-specific outcomes, 10 focused on students with literacy challenges (of these, 8 were conducted in elementary grades), and reported effect sizes ranging from -0.03 to 0.92, with most effect sizes greater than 0.45. Balanced instructional programs with reading outcomes for students with literacy challenges included the following: Remedial instruction (explicit instruction and practice on skills and processes essential to literacy development;  $n = 6$ ,  $ES = -0.03$  to  $0.54$ ), Writing to Read (computer-based instruction focusing on phonics, listening skills, reading along while listening to books, and typing stories;  $n = 2$ ,  $ES = 0.45$  to  $0.92$ ), cooperative learning (cooperative learning activities embedded in explicit reading instruction integrated with a process approach to writing;  $n = 1$ ,  $ES = 0.66$ ), and whole language plus early literacy (instruction emphasizing meaning-making and a process approach to writing along with instruction on emerging literacy skills such as phonological awareness and word reading;  $n = 1$ ,  $ES = 0.86$ ).

We were unable to discern specific effects on word-reading outcomes for students with literacy challenges because authors did not break down these effect sizes by type of reading measure. Consistent with previous syntheses of the effects of writing instruction on reading, these findings suggest that instruction that includes a balance of reading and writing instruction can have a positive effect on reading, including for students who struggle.

### ***Summary of Effects of Writing Instruction and Intervention***

Overall, findings from these three meta-analyses indicate that instruction and intervention focused on writing skills or a balance of reading and writing skills can have a positive impact on reading outcomes for students with and without reading and writing difficulties. This instruction included word-level focused instruction (e.g., spelling) as well as sentence- and passage-level focused instruction (e.g., sentence construction, writing process), reinforcing the “shared

knowledge” perspective that supporting development in multiple dimensions of writing can support development in reading. These findings suggest that writing might be an important dimension of literacy instruction designed to improve reading outcomes.

### **Discussion**

The purpose of this synthesis of meta-analyses and systematic reviews was to determine what we know and still need to learn about effective literacy instruction for students with reading difficulties and disabilities, including dyslexia. We examined reviews conducted over the past decade that addressed our three research questions about effects of: (a) early literacy interventions in the early elementary grades, (b) reading interventions conducted in the later elementary grades, and (c) writing instruction and interventions on reading skills. Findings across reviews revealed some clear directions for practice (what we know) and future research (what we need to know) focused on improving outcomes for children with reading difficulties and disabilities.

#### **What We Know**

##### ***Primary Grade Interventions***

Based on findings from six syntheses of studies conducted within the primary grades, we know there is converging causal evidence for the effectiveness of reading interventions to guide instructional practices for special educators, reading interventionists, and classroom teachers. Notably, as summarized in text and in Table 1, several syntheses reported a consistent and statistically significant overall mean effect size of 0.39. Across these syntheses and meta-analyses, the magnitude of effect sizes on standardized reading measures was stronger on code-focused measures (ranging from 0.41 to 0.62) than on meaning-focused measures (ranging from 0.32 to 0.36). Thus, intervention research has not yet revealed findings for reading

comprehension that are as robust, although this is the ultimate goal of reading. This finding is consistent with the *Simple View of Reading* (Gough & Tunmer, 1986), suggesting that if readers' primary difficulty is word-level reading, even if word reading skill is improved, it is also important to support comprehension.

None of the moderator analyses indicated differential effects based on student risk which indicates that, on average, a variety of struggling readers, including students with word-level reading difficulties and those with reading disabilities, benefited from the explicit and systematic interventions studied. Austin et al. (2017) specifically focused on students who had demonstrated a lack of response to high quality instruction and reported moderate to large effects for intensive interventions relative to control conditions in a small number of studies, indicating promise for interventions provided within multi-tiered systems of support.

We also have some knowledge about important intervention features to inform research and guide practice. Most interventions were delivered in small groups or one-to-one and provided explicit and systematic code- and meaning-focused instruction, consistent with recommendations for multi-tiered systems of support and response to intervention made by Gersten et al. (2009). When there were enough studies to conduct moderator analyses or meta-regression, findings revealed no differences in effectiveness of interventions based on group size, the dosage of intervention, or the type of interventionists. However, Slavin and colleagues (2011), who did not conduct a formal moderation analysis, did report one-to-one tutoring was the most robust and should be considered a gold standard. They also acknowledged this type of tutoring is costly and should be reserved for students with the greatest need. Further, they noted that interventions provided by teachers had stronger effects than paraprofessionals, and the weakest effects were found when volunteers were implementing. Within moderation analyses,

the type of intervention did not seem to moderate effects across the primary grades, with one exception; Gersten et al. (2020) noted that when phonics interventions included spelling, they were more effective than when they included phonemic awareness. In addition, Slavin et al. (2011) reported interventions that followed an explicit and systematic approach had stronger effects than those that did not. Al Otaiba et al. (2018) found no evidence of stronger effects for dyslexia-specific approaches that incorporated multi-sensory approaches relative to other explicit and systematic approaches, but this is based on only one study. Similarly, Stevens et al. (2021) examined a broader corpus of research, and found no significant differences favoring OG multisensory approaches relative to typical instruction. This meta-analysis included participants in the primary and upper elementary grades and a small number of studies with older students, and the authors but did not explore moderation of effects by age or grade.

### *Upper Elementary Interventions*

In addition, findings from the four meta-analyses that included upper elementary students provide causal evidence for the efficacy of reading interventions for students with reading difficulties or disabilities in these grades, albeit with smaller effects. Several meta-analyses revealed smaller effects on standardized measures than on researcher-made measures (Donegan & Wanzek, 2021; Flynn et al., 2012; Scammacca et al., 2013). This pattern was consistent for foundational code-focused skills and for comprehension (Donegan & Wanzek, 2021). However, Flynn et al. (2012) reported an overall mean effect size of 0.37, which is more consistent with the findings of studies in the primary grades. This meta-analysis reported the largest effect sizes among those reviewed on standardized comprehension measures (0.73), but also moderate and negative effect sizes on reading fluency measures (-0.29). Notably, none of these syntheses found that age or grade moderated effects. Thus, across these four syntheses, on average, a

variety of struggling readers, including upper elementary students at-risk for or with reading disabilities benefited from explicit and systematic interventions. None of the reviews specifically described response to previous intervention of the participants, although they had clearly not benefited sufficiently from typical school instruction in the earlier grades to read on grade level.

We also know about some promising features related to intervention intensity for upper elementary students based on moderator analyses summarized in Table 1. For example, in one meta-analysis (Scammacca et al., 2013), comprehension interventions had a higher effect for students than did other types of interventions. Yet, another meta-analysis (Flynn et al., 2012) revealed no significant moderators, including instructional focus, but authors noted that studies with moderate positive effect sizes shared intervention features such as explicit and systematic instruction. These analyses included between zero (Flynn et al., 2012) and three (Scammacca et al., 2013) studies that were specific to students in the upper elementary grades.

However, examination of the research specific to upper elementary students did provide different findings on moderation of outcomes that may be helpful for instruction. Although the syntheses that included Grade 4-12 did not find moderation by group size, duration, or instructional component, Donegan and Wanzek (2021) did note some significant moderation for the research specific to upper elementary students. Their findings suggest that multicomponent interventions may be the most helpful for elementary students with poor reading achievement after Grade 3. Additionally, small group instruction provided students with the greatest chance of improving both their foundational and comprehension outcomes. The smallest groups (1-2 students) may be needed to better affect comprehension outcomes in particular. Similarly,

affecting comprehension outcomes for these upper elementary students with reading difficulties or disabilities may require longer duration interventions.

Notably, individualizing interventions was not as effective as more standardized interventions for improving foundational or comprehension outcomes. Of course, the standardized interventions did not ignore student progress, with students progressing through lessons only as they mastered previous lessons, but the general structure of the lessons and what was taught at each step in the student's learning was standardized across students and pre-planned in the lessons rather than the teacher determining the lesson at each step for each student. The finding may be an effect of the difficulty with accurately individualizing instruction for different students or the difficulty associated with teachers being tasked with creating accelerated instruction along the way for multiple students. It is also possible that the group size provided easier opportunities for teachers to individualize within standardized lessons than creating those lessons for each student. Thus, generally speaking, the upper elementary research suggests small, positive effects can be achieved when standardized, small-group instruction addressing multiple components are provided to students with reading difficulties or disabilities who have low reading achievement in foundational outcomes. Comprehension outcomes may benefit from longer duration interventions.

### ***Writing Instruction to Improve Reading Outcomes***

In addition to substantial evidence of the effectiveness of explicit, systematic intervention for students who experience significant reading difficulties across the elementary grades, there is promising evidence that writing instruction and interventions can have an important impact on reading outcomes. Three meta-analyses with this focus provide consistently positive support for this conclusion for students overall, including preliminary evidence for those who experience

literacy-related challenges. Graham and Hebert (2011) revealed moderate effects of transcription and sentence-level instruction on word-reading (0.39 to 0.51) and fluency (0.32) outcomes and relatively weaker effects for text-generation instruction on reading comprehension outcomes (0.17 to 0.30) for students identified with reading and writing difficulties. Similarly, Graham and Santangelo (2014) found moderate effects of explicit spelling instruction on phonological awareness (0.51) for students in primary grades, as well as on word-reading, fluency, and general measures of reading for students with literacy difficulties in Grades K-6 (with most effect sizes greater than 0.30). Further, Graham et al. (2018) provided evidence of strong, positive effects of balanced reading and writing instruction for students overall as well as those experiencing literacy challenges, though ES varied widely (-0.03 to 0.92, with most ES greater than 0.45). Moderator analyses suggest that balanced reading and writing instruction may be more effective for older students and when it is equally balanced rather than heavier on reading; however, these findings pertained to students across the full range of literacy performance and thus firm conclusions cannot be drawn for those experiencing literacy difficulties.

### **What We Still Need to Learn**

Findings from this review of reviews should be interpreted in the context of a few limitations. We intentionally limited our synthesis to meta-analyses and systematic reviews conducted over the past decade to reflect the time frame when legislation in many states have mandated dyslexia-specific interventions. We could have done a more traditional systematic review or meta-analysis, but given the existence of multiple meta-analyses, our ‘review of reviews’ approach allowed us to address our questions more broadly and comprehensively. By virtue of synthesizing meta-analyses and systematic reviews that only included experimental and quasi-experimental studies, or single case designs, we may have overlooked important research

using other designs, which could shed additional light on what we know about effective literacy intervention for students with dyslexia, and students with different types of reading difficulties.

As within any review of reviews, it can be challenging to provide a broad survey of what works, while still focusing on identifying key instructional components. Until the recent review focused specifically on upper grades by Donegan and Wanzek (2021), it was a challenge to derive conclusions for upper elementary students. Therefore, we included some meta-analyses that included studies with students in the upper elementary grades even if they also spanned into middle or high school. Other challenges in interpretation related to differences in how participants were identified for studies (e.g., at risk, with reading difficulties or disabilities, rather than only dyslexia). Furthermore, we were only able to include studies of writing instruction broadly rather than focusing on intervention for students with specific difficulties or disabilities in this area (given the lack of relevant meta-analyses).

Beyond these limitations, a number of questions remain unanswered in the current literature. Several meta-analyses included moderator analyses related to features that relate to intensity, but generally, effects did not vary by student risk status, intervention type, grade level, implementer training, group size, or hours of intervention in the primary grades. The consistent lack of significant moderators in the existing literature in the early elementary grades should not lead us to conclude that none of these features of intensity matter; rather, there is simply not enough evidence in the current literature to draw firm conclusions. Interestingly, none of the moderator analyses included gender, ethnicity or race, or socioeconomic status. This omission is problematic in light of recent research indicating that minoritized students are less likely than white students to receive a school-provided dyslexia classification (and services) as are students attending schools with higher proportions of students with poor literacy skills (i.e., Odegard et

al., 2020). This body of research also did not explore whether the effects varied for students with a family history of dyslexia, or who experience reading difficulties that are in specific areas or multiple areas (e.g., word reading-only or word reading and comprehension). There were also relatively few rigorous studies examining dyslexia-specific interventions (Stevens et al., 2021). This lack of evidence is concerning in light of legislation that may mandate dyslexia-specific intervention, whether it is integrated into MTSS or whether it is implemented on top of existing efforts because such additional intervention requirements may further tax under-resourced schools that already struggle to implement effective instruction and intervention. Thus, we still need to know which features optimize effectiveness of intensive intervention (e.g., Fuchs et al., 2014; Wanzek et al., 2020). Relatedly, at the early elementary level, a majority of studies were conducted in kindergarten and first grade, when early literacy interventions are intended to prevent or ameliorate risk for reading disabilities, including dyslexia. Within the primary grades, additional research with students in Grades 2-3 will help clarify the essential elements of intensive intervention for students who have not responded to prior early intervention. At the upper elementary level, important features of intervention did emerge from moderator analyses; however, more evidence is also needed to better understand for whom (e.g., students with dyslexia vs. other reading difficulties/disabilities) and under what conditions (e.g., multi-component vs. more targeted) interventions will work best.

Given that many of the moderators explored across the meta-analyses we reviewed were not significant at either primary or upper elementary grades, future research studies are needed that include larger number of students with different student characteristics and reading profiles. It may also be possible to combine smaller extant data samples into larger data sets using data

fusion with integrative data analytic methods (e.g., Daucourt et al., 2018), Researchers should also include pre-treatment word reading scores more consistently.

Along these lines, across grade levels, there is a need for more studies that identify ways to match student need to intervention (e.g., Fuchs et al., 2014). Several syntheses have identified predictors of responsiveness to early reading intervention (e.g., Al Otaiba & Fuchs, 2002; Lam & McMaster, 2014; Nelson et al., 2003; Stuebing et al., 2015; Tran et al., 2011), which may be informative in determining what works and what does not within approaches to intensifying intervention for students with the most significant learning needs. Slavin et al. (2011) concluded that 1:1 tutoring provided by expert teachers should be considered a gold standard, and in their best evidence synthesis they produced higher effects than did paraprofessionals or volunteers. However, in most other meta-analyses that examined implementer as a moderator, this pattern did not appear; rather, a variety of implementers provided intervention effectively. Future research is needed to test the effects of implementer expertise, and whether the need for expertise and training interacts with the complexity of interventions or with the level of student need. Certainly, we need to test the efficacy of dyslexia-specific interventions, including those that are multi-sensory, given strong advocacy for these approaches by dyslexia organizations, state legislation, and policy (e.g., [improvingliteracy.org/state-of-dyslexia](http://improvingliteracy.org/state-of-dyslexia)) - and for whom such interventions are likely to be most effective. For example, Stevens et al. (2021) reported no significant difference in reading outcomes for Orton-Gillingham-based programs relative to control groups for any foundational skills including phonological awareness, phonics, fluency or spelling, although average effect sizes were positive (mean ES = 0.22). As noted by Ritchey & Goeke (2006), Al Otaiba et al. (2018) and most recently by Stevens et al. (2021), future research

should include randomized control trials comparing these approaches to other synthetic phonics approaches.

Further research is also needed to evaluate differential effects of standard treatment protocols vs. interventions individualized based on data (c.f., Wanzek et al., 2018). In fact, researchers have provided promising evidence for using data-based individualization to intensify and individualize instruction for students with significant learning needs (e.g., Connor et al., 2013). An associated direction for future research relates to discovering how to enhance systematic phonics interventions for students with word-level reading and spelling difficulties. For example, researchers have examined individual differences in students' mental "set for variability" (e.g., Venezky, 1999), or their awareness that not all grapheme phoneme correspondences are regular in English (Kearns et al., 2016). Some evidence suggests it is effective to train students to be more flexible decoders particularly when reading words with variable vowels or multi-syllabic words (e.g., Lovett et al., 2014; Savage et al., 2018; Steacy et al., 2016; Steacy et al., 2019; Tunmer & Chapman, 2012). Students' vocabulary knowledge and their phonemic awareness may be important avenues to facilitate flexibility (e.g., Nation & Castles, 2017; Share, 1995; Steacy et al., 2016). Relatedly, we need to learn much more about the types of spelling and writing instruction that are most beneficial for students with reading difficulties, and how best to balance writing and reading intervention to maximize outcomes in both areas (Graham, 2020; Graham & Santangelo, 2014).

In line with all of the above points, as a field we need to better operationalize, and perhaps come to consensus on, how we define/identify students with 'dyslexia,' including what constitutes 'nonresponsiveness' to intervention (e.g., Austin et al., 2017). Definitions of dyslexia, reading difficulty, or reading disability (see Elliott, 2020) may ultimately lead to differences in

how those participants are identified as eligible for, and how they respond to different types of interventions. Currently, limited information in the majority of studies on subtypes of reading disability or difficulty, including whether students are identified with word reading difficulties, comprehension difficulties, or both also make it difficult to interpret whether reading interventions specifically address one or more of these subtypes of disability. Overall, converging evidence suggests some universal features of intervention that should be considered; on the other hand, more consistency in identification of students with dyslexia would firm up our understanding of whether specific interventions are differentially effective for this group.

### **Conclusion**

This review of reviews sheds important light on what we know and need to learn about effective literacy interventions for elementary children who experience significant difficulty learning to read, and specifically those with word-level reading difficulties that ultimately affect their ability to comprehend and learn from text. First, we know that explicit and systematic intensive intervention, focusing on the code *and* meaning dimensions of reading, and delivered one-to-one or in small groups, are likely to improve foundational code-based reading skills, and to a lesser extent, meaning-based skills, across grade levels. We know that, at least in the upper elementary grades, some intervention features including standardized protocols, multiple components, and longer duration yields stronger effects. And we know that balancing reading and writing intervention shows promise. We still have much to learn regarding *for whom* and *under what conditions* literacy interventions work best. Further research examining individual differences and how best to individualize interventions for students identified as having dyslexia or other reading difficulties and disabilities should continue to advance our understanding of how best to meet these students' significant learning needs.

### References

- Adams, M. J. (1990). Beginning to read: Thinking and learning about print. *About Print*.
- Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention. *Remedial and Special Education, 23*(5), 300-316.  
<https://doi.org/10.1177/07419325020230050501>
- Al Otaiba, S., Rouse, A. G., & Baker, K. (2018). Elementary grade intervention approaches to treat specific learning disabilities, including dyslexia. *Language, Speech & Hearing Services in Schools, 49*(4), 829-842.  
[https://doi.org/10.1044/2018\\_LSHSS-DYSLC-18-0022](https://doi.org/10.1044/2018_LSHSS-DYSLC-18-0022)
- Austin, C. R., Vaughn, S., & McClelland, A. M. (2017). Intensive reading interventions for inadequate responders in grades k–3: a synthesis. *Learning Disability Quarterly, 40*(4), 191-210. <https://doi.org/10.1177/0731948717714446>
- Berninger, V. W., Abbott, R. D., Abbott, S. P., Graham, S., & Richards, T. (2002). Writing and reading. *Journal of Learning Disabilities, 35*(1), 39-56.  
<https://doi.org/10.1177/002221940203500104>
- Birsch, J.R. (Ed.). (2011). *Multisensory teaching of basic language skills* (3rd ed.). Baltimore, MD: Paul H. Brookes.
- Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: reading acquisition from novice to expert. *Psychological Science in the Public Interest, 19*(1), 5-51.  
<https://doi.org/10.1177/1529100618772271>
- Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49*(2), 278-293. [https://doi.org/10.1044/1092-4388\(2006/023\)](https://doi.org/10.1044/1092-4388(2006/023))

- Cervetti, G. N., Pearson, P. D., Palincsar, A. S., Afflerbach, P., Kendeou, P., Biancarosa, G., ... & Berman, A. I. (2020). How the reading for understanding initiative's research complicates the simple view of reading invoked in the science of reading. *Reading Research Quarterly*, 55, S161-S172. <https://doi.org/10.1002/rrq.343>
- Cirino, P. T., Romain, M. A., Barth, A. E., Tolar, T. D., Fletcher, J. M., & Vaughn, S. (2012). Reading skill components and impairments in middle school struggling readers. *Reading & Writing*, 26(7), 1059-1086. <https://doi.org/10.1007/s11145-012-9406-3>
- Connor, C. M., Morrison, F. J., Fishman, B., Crowe, E. C., Al Otaiba, S., & Schatschneider, C. (2013). A longitudinal cluster-randomized controlled study on the accumulating effects of individualized literacy instruction on students' reading from first through third grade. *Psychological Science*, 24, 1408-1419. PMID: PMC4737583
- Conrad, N. J. (2008). From reading to spelling and spelling to reading: Transfer goes both ways. *Journal of Educational Psychology*, 100(4), 869.
- Daucourt, M. C., Schatschneider, C., Connor, C. M., Al Otaiba, S., & Hart, S. A. (2018). Inhibition, updating working memory, and shifting predict reading disability symptoms in a hybrid model: Project KIDS. *Frontiers in Psychology*, 9, 238. <https://doi.org/10.3389/fpsyg.2018.00238>
- Denton, C. A., Fletcher, J. M., Anthony, J. L., & Francis, D. J. (2006). An evaluation of intensive intervention for students with persistent reading difficulties. *Journal of Learning Disabilities*, 39(5), 447-466. <https://doi.org/10.1177/00222194060390050601>
- Denton, C. A., Tolar, T. D., Fletcher, J. M., Barth, A. E., Vaughn, S., & Francis, D. J. (2013).

- Effects of tier 3 intervention for students with persistent reading difficulties and characteristics of inadequate responders. *Journal of Educational Psychology*, 105(3), 633-648. <https://doi.org/10.1037/a0032581>
- Donegan, R. E., & Wanzek, J. (2021). Effects of reading interventions implemented for upper elementary struggling readers: A look at recent research. *Reading and Writing*, 1-35. <https://doi.org/10.1007/s11145-021-10123-y>
- Ehri, L. C. (1987). Learning to read and spell words. *Journal of reading behavior*, 19(1), 5-31.
- Ehri, L. C. (2000). Learning to read and learning to spell: two sides of a coin. *Topics in Language Disorders*, 20(3), 19-36. <https://doi.org/10.1097/00011363-200020030-00005>
- Ehri, L. C. (2005). Learning to read words: theory, findings, and issues. *Scientific Studies of Reading*, 9(2), 167–188. [https://doi.org/10.1207/s1532799xssr0902\\_4](https://doi.org/10.1207/s1532799xssr0902_4)
- Ehri, L. C. (2020). The science of learning to read words: a case for systematic phonics instruction. *Reading Research Quarterly*, 55(S1), S45-S60. <https://doi.org/10.1002/rrq.334>
- Elliott, J. G. (2020). It's time to be scientific about dyslexia. *Reading Research Quarterly*, 55(S1), S61-S75. <https://doi.org/10.1002/rrq.333>
- Elliott, J. G., & Grigorenko, E. L. (2014). *The dyslexia debate* (No. 14). Cambridge University Press.
- Every Student Succeeds Act. (2015). Title 20 § 6301 to 7981. Retrieved from: <http://www.gpo.gov/fdsys/pkg/BILLS-114s1177enr/pdf/BILLS-114s1177enr.pdf>
- Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist*, 35(1), 39-50. [https://doi.org/10.1207/s15326985ep3501\\_5](https://doi.org/10.1207/s15326985ep3501_5)
- Flynn, L. J., Zheng, X., & Swanson, H. L. (2012). Instructing struggling older readers: A

- selective meta-analysis of intervention research. *Learning Disabilities Research & Practice*, 27(1), 21–32. <https://doi.org/10.1111/j.1540-5826.2011.00347.x>
- Fuchs, D., Fuchs, L. S., & Vaughn, S. (2014). What is intensive instruction and why is it important? *Teaching Exceptional Children*, 46(4), 13-18.  
<https://doi.org/10.1177/0040059914522966>
- Gersten, R., Compton, D., Connor, C.M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W.D. (2009). Assisting students struggling with reading: response to intervention (RtI) and multi-tier intervention in the primary grades (NCEE 2009–4045). Washington, DC: *National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences*, U.S. Department of Education.
- Gersten, R., Haymond, K., Newman-Gonchar, R., Dimino, J., & Jayanthi, M. (2020). Meta-analysis of the impact of reading interventions for students in the primary grades. *Journal of Research on Educational Effectiveness*, 13(2), 401-427.  
<https://doi.org/10.1080/19345747.2019.1689591>
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *RASE: Remedial & Special Education*, 7(1), 6–10. <https://doi.org/10.1177/074193258600700104>
- Graham, S. (2020). The sciences of reading and writing must become more fully integrated. *Reading Research Quarterly*, 55(S1), S35-S44. <https://doi.org/10.1002/rrq.332>
- Graham, S., & Hebert, M. (2011). Writing to read: a meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, 81(4), 710-744.  
<https://doi.org/10.17763/haer.81.4.t2k0m13756113566>
- Graham, S., Liu, X., Aitken, A., Ng, C., Bartlett, B., Harris, K. R., & Holzapfel, J. (2018).

- Effectiveness of literacy programs balancing reading and writing instruction: a meta-analysis. *Reading Research Quarterly*, 53(3), 279-304. <https://doi.org/10.1002/rrq.194>
- Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading & Writing*, 27(9), 1703-1743. <https://doi.org/10.1007/s11145-014-9517-0>
- Grigorenko, E. L., Compton, D. L., Fuchs, L. S., Wagner, R. K., Willcutt, E. G., & Fletcher, J. M. (2020). Understanding, educating, and supporting children with specific learning disabilities: 50 years of science and practice. *The American Psychologist*, 75(1), 37–51. <https://doi.org/10.1037/amp0000452>
- Gunn, B. K. (1996). An investigation of three approaches to teaching phonological awareness to first-grade students and the effects on word recognition (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9706736)
- Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., and Torgesen, J. (2008). Improving adolescent literacy: effective classroom and intervention practices: a practice guide (NCEE #2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, *Institute of Education Sciences, U.S. Department of Education*. Retrieved from <http://ies.ed.gov/ncee/wwc>.
- Kearns, D. M., Rogers, H. J., Koriakin, T., & Al Ghanem, R. (2016). Semantic and phonological ability to adjust recoding: A unique correlate of word reading skill. *Scientific Studies of Reading*, 20(6), 455-470. <https://doi.org/10.1080/10888438.2016.1217865>
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6(2), 293-323.

[https://doi.org/10.1016/0010-0285\(74\)90015-2](https://doi.org/10.1016/0010-0285(74)90015-2)

- Lam, E. A. & McMaster, K. L. (2014). Predictors of responsiveness to early literacy intervention: A 10-Year Update. *Learning Disability Quarterly*, 37(3), 134-147.  
<https://doi.org/10.1177/0731948714529772>
- Lindamood, P. & Lindamood, P. (1998). The Lindamood phoneme sequencing program for reading, spelling, and speech. Austin: PRO-ED.
- Lovett, M. W., Lacerenza, L., Steinbach, K. A., & De Palma, M. (2014). Development and evaluation of a research-based intervention program for children and adolescents with reading disabilities. *Perspectives on Language and Literacy*, 40(3), 21-31.  
<https://doi.org/10.1016/j.tate.2007.10.005>
- Miciak, J., & Fletcher, J. M. (2020). The critical role of instructional response for identifying dyslexia and other learning disabilities. *Journal of Learning Disabilities*, 53(5), 343-353.  
<https://doi.org/10.1177/0022219420906801>
- Moats, L. C. (2005). How spelling supports reading. *American Educator*, 6(12-22), 42.
- National Center for Education Statistics. (2019). *NAEP: The nation's report cards. An overview of NAEP*.
- National Center on Improving Literacy. (2020). State of Dyslexia. Retrieved from  
<https://improvingliteracy.org/state-of-dyslexia>
- National Reading Panel (US). (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.
- Nation, K., & Castles, A. (2017). Putting the learning into orthographic learning. *Theories of*

*reading development*, 15, 147-168.

Nelson, J., Benner, G. J., & Gonzalez, J. (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review. *Learning Disabilities Research and Practice*, 18(4), 255-267. <https://doi.org/10.1111/1540-5826.00080>

Odegard, T. N., Farris, E. A., Middleton, A. E., Oslund, E., & Rimrodt-Frierson, S. (2020). Characteristics of students identified with dyslexia within the context of state legislation. *Journal of Learning Disabilities*, 53(5), 366-379. <https://doi.org/10.1177/0022219420914551>

Orton, S. T. (1937). *Reading, writing and speech problems in children*. Norton.

Petscher, Y., Cabell, S.Q., Catts, H.W., Compton, D.L., Foorman, B.R., Hart, S.A., Lonigan, C.J., Phillips, B.M., Schatschneider, C., Steacy, L.M., Terry, N.P., & Wagner, R.K. (2020). How the science of reading informs 21st-Century education. *Reading Research Quarterly*, 55(S1), S267– S282. <https://doi.org/10.1002/rrq.352>

Pinnell, G., DeFord, D., & Lyons, C. (1988) *Reading Recovery: early intervention for at-risk first graders*. Arlington:ERS.

Ritchey, K. D., & Goeke, J. L. (2006). Orton-Gillingham and Orton-Gillingham—based reading instruction. *The Journal of Special Education*, 40(3), 171-183. <https://doi.org/10.1177/00224669060400030501>

Samuels, S. J., Ediger, K. A., & Fautsch-Patridge, T. (2005). The importance of fluent reading. *New England Reading Association Journal*, 41(1), 1.

Savage, R., Georgiou, G., Parrila, R., & Maiorino, K. (2018). Preventative reading interventions

teaching direct mapping of graphemes in Texts and set-for-variability aid at-risk learners. *Scientific Studies of Reading*, 22(3), 225–247.

<https://doi.org/10.1080/10888438.2018.1427753>

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2013). A meta-analysis of interventions for struggling readers in grades 4–12. *Journal of Learning Disabilities*, 48(4), 369-390. <https://doi.org/10.1177/0022219413504995>

Scanlon, D. M., Vellutino, F. R., Small, S. G., Fanuele, D. P., & Sweeney, J. M. (2005). Severe reading difficulties-can they be prevented? A comparison of prevention and intervention approaches. *Exceptionality*, 13(4), 209-227. [https://doi.org/10.1207/s15327035ex1304\\_3](https://doi.org/10.1207/s15327035ex1304_3)

Schlesinger, N., & Gray, S. (2017). The impact of multisensory instruction on learning letter names and sounds, word reading, and spelling. *Annals of Dyslexia*, 67(3), 219-258.

<https://doi.org/10.1007/s11881-017-0140-z>

Seidenberg, M. S., Cooper Borkenhagen, M., & Kearns, D. M. (2020). Lost in translation? Challenges in connecting reading science and educational practice. *Reading Research Quarterly*, 55. <https://doi.org/10.1002/rrq.341>

Shanahan, T. (2016). Thinking with research: research changes its mind (again). *The Reading Teacher*, 70(2), 245-248. <https://doi.org/10.1002/trtr.1504>

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, 55(2), 151-218.

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, 6(1), 1-26.

<https://doi.org/10.1016/j.edurev.2010.07.002>

Snow, C., Biancarosa, G. (2003) Adolescent literacy and the achievement gap: what do we

know and where do we go from here? *Carnegie Corporation of New York*.

Snowling, M. J., Hulme, C., & Nation, K. (2020). Defining and understanding dyslexia: past, present and future. *Oxford Review of Education*, 46(4), 501-513.

<https://doi.org/10.1080/03054985.2020.1765756>

Steady, L.M., Compton, D.L., Petscher, Y., Elliott, J.D., Smith, K., Rueckl, J.G., Sawi, O., Frost, S.J., & Pugh, K.R. (2019). Development and prediction of context-dependent vowel pronunciation in elementary readers, *Scientific Studies of Reading*, 23:1, 49-63.

<https://doi.org/10.1080/10888438.2019.1620749>

Steady, L. M., Elleman, A. M., Lovett, M. W., & Compton, D. L. (2016). Exploring differential effects across two decoding treatments on item-level transfer in children with significant word reading difficulties: A new approach for testing intervention elements.

*Scientific Studies of Reading*, 20(4), 283-295.

<https://doi.org/10.1080/10888438.2016.1178267>

Stevens, E. A., Austin, C., Moore, C., Scammacca, N., Boucher, A. N., & Vaughn, S. (2021).

Current state of the evidence: Examining the effects of Orton-Gillingham reading interventions for students with or at risk for word-level reading disabilities. *Exceptional Children*, <https://doi.org/0014402921993406>.

Stuebing, K., Barth, A., Trahan, L., Reddy, R., Miciak, J., & Fletcher, J. (2015). Are child cognitive characteristics strong predictors of responses to intervention? A meta-analysis.

*Review of Educational Research*, 85(3), 395-429.

<https://doi.org/10.3102/0034654314555996>

Treiman, R. (1993). *Beginning to spell: A study of first-grade children*. Oxford University Press on Demand.

- Torgesen, J., Myers, D., Schirm, A., Stuart, E., Vartivarian, S., Mansfield, W., et al. (2006). National assessment of title I interim report: volume II: closing the reading gap: first year findings from a randomized trial of four reading interventions for striving readers. Washington, DC: U.S. Department of Education, *Institute of Education Sciences*.
- Torgesen, J., Wagner, R., Rashotte, C., Herron, J., & Lindamood, P. (2010). Computer-assisted instruction to prevent early reading difficulties in students at risk for dyslexia: outcomes from two instructional approaches. *Annals of Dyslexia*, 60(1), 40-56.  
<https://doi.org/10.1007/s11881-009-0032-y>
- Tran, L., Sanchez, T., Arellano, B., & Lee Swanson, H. (2011). A meta-analysis of the RTI literature for children at risk for reading disabilities. *Journal of Learning Disabilities*, 44(3), 283-295. <https://doi.org/10.1177/0022219410378447>
- Tunmer, W. E., & Chapman, J. W. (2012). The simple view of reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities*, 45(5), 453-466. <https://doi.org/10.1177/0022219411432685>
- Vaughn, S., Linan-Thompson, S., & Hickman, P. (2003). Response to instruction as a means of identifying students with reading/learning disabilities. *Exceptional Children*, 69(4), 391-409. <https://doi.org/10.1177/001440290306900401>
- Venezky, R. L. (1999). *The American way of spelling: The structure and origins of American English orthography*. Guilford Press.
- Wanzek, J., Otaiba, S. A., & McMaster, K. L. (2020). *Intensive Reading Interventions for the Elementary Grades*. The Guilford Press.
- Wanzek, J., Stevens, E. A., Williams, K. J., Scammacca, N., Vaughn, S., & Sargent, K. (2018).

- Current evidence on the effects of intensive early reading interventions. *Journal of Learning Disabilities*, 51(6), 612-624. <https://doi.org/10.1177/0022219418775110>
- Wanzek, J., Vaughn, S., Scammacca, N., Gatlin, B., Walker, M. A., & Capin, P. (2015). Meta-analyses of the effects of tier 2 type reading interventions in grades K-3. *Educational Psychology Review*, 28(3), 551-576. <https://doi.org/10.1007/s10648-015-9321-7>
- Wanzek, J., Vaughn, S., Scammacca, N. K., Metz, K., Murray, C. S., Roberts, G., & Danielson, L. (2013). Extensive reading interventions for students with reading difficulties after grade 3. *Review of Educational Research*, 83(2), 163-195. <https://doi.org/10.3102/0034654313477212>
- What Works Clearinghouse (2010<sup>a</sup>). Unbranded Orton-Gillingham-based interventions. Washington, DC: U.S. Department of Education, Institute of Education Sciences. *Mathematica Policy Research*. <https://ies.ed.gov/ncee/wwc/EvidenceSnapshot/528>
- What Works Clearinghouse (2010<sup>b</sup>). Wilson Reading. Washington, DC: U.S. Department of Education, Institute of Education Sciences. [https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc\\_wilson\\_070110.pdf](https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_wilson_070110.pdf)
- What Works Clearinghouse (2010<sup>c</sup>). Lindamood Phoneme Sequencing. Washington, DC: U.S. Department of Education, Institute of Education Sciences. [https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc\\_lindamood\\_111015.pdf](https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_lindamood_111015.pdf)
- What Works Clearinghouse (2013). Procedures and standards handbook, Version 3.0. Washington, DC: U.S. Department of Education, Institute of Education Sciences. [https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc\\_procedures\\_v3\\_0\\_standards\\_handbook.pdf](https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf)

Wilson, B. 1988. Wilson Reading System. Millbury, MA: Wilson Language Training.

Youman, M. & Mather, N. (2018). Dyslexia laws in the USA: a 2018 update. *Perspectives on Language and Literacy, Spring*, 37-41.

**Table 1***Summaries of Syntheses and Main Findings for Reading Outcomes*

Study	Research design	Grade level	Risk definition	Dosage	Number of studies	Mean ES	Moderators
Effects of reading interventions in grades K-3							
Slavin et al. (2011)	E, QED	K - 5	SWRD (S < 33% classroom reading performance; receiving intensive reading intervention)	12 weeks	<i>n</i> = 20 (teacher 1x1)	0.39	NR
					<i>n</i> = 18 (Para/Vol 1x1)	0.24	
					<i>n</i> = 20 (small group)	0.31	
Wanzek et al. (2015)	E, QED	K - 3	SWRD, at-risk or struggling readers	15-99 sessions	72	0.54 CF 0.36 MF	Intervention type, Grade level, Implementer, Group size, Hours of intervention
Wanzek et al. (2018)	E, QED	K - 3	SWLD, at-risk (low ability, low PA, language disorders)	100 sessions	25	0.39 overall	Intervention type, Grade Level, Implementer, Group size, Hours of intervention

Gersten et al. (2020)		1 - 3	Risk = < 25%ile; Min risk = < 40th%ile	8 hours	33	0.39 Overall 0.41 Word reading 0.31 Fluency 0.32 Comp.	Intervention type (focus: PA, Spelling, Writing)*, Grade level, Implementer, Group size, Hours of intervention Student risk, Type of comparison, Design, Measurement domain,* Outcome measure
Stevens et al. (2021)	E, QED, SCD	K-12	Dyslexia, LD, reading disabilities/difficulty; low performance on standardized measure	NA	16	0.22 Found. 0.14 Vocab & Comp.	Research design, Comparison condition, Implementation fidelity, Publication year*

Effects of reading interventions in grades 4-12

Flynn et al. (2012)	E, QED	5 - 9	Reading disabilities (scores below 25th%ile on norm-referenced measure)	5-9 sessions	11	NR	Intervention type, Number of sessions, Length of session, Weeks of intervention, Age, Measure Type, Grade
Wanzek et al. (2013)	E, QED	4 - 12	SWLD, reading difficulties (reading below grade level)	75 sessions	10	0.16 Fluency 0.15 Word recog. 0.16 Word fluency 0.15 Spelling 0.10 Comp.	Hours of intervention Group size Grade
Scammacca et al. (2013)	E, QED	4 - 12	Struggling readers (low achievement in reading, unidentified reading difficulties, dyslexia, and/or with reading or LD)	NR	12	0.49	Intervention type <sup>ab</sup> , Grade, Implementer <sup>b</sup> , LD status <sup>b</sup> , Hours of intervention, Type of comparison

Donegan et al. (2021)	E, QED	4 - 5	Reading disabilities, at-risk for reading disabilities, or below grade level in reading	15 sessions	33	0.22 Found. 0.21 Comp.	Intervention Type <sup>cd</sup> Group size <sup>cd</sup> , Hours of intervention <sup>d</sup> , Standardization <sup>cd</sup>
Effects of writing instruction in grades PreK-12							
Graham et al. (2011)	E, QED	1 - 12	All levels of reading or writing achievement	NA	21	0.22 Comp.	NA
Graham et al. (2014)	E, QED	K - 12	All levels of reading or writing achievement	NA	27	0.51 PA 0.44 Reading	Quantity of Spelling, Instruction
Graham et al. (2018)	E, QED	PreK - 12	All levels of reading or writing achievement	NA	38	0.53 Word reading 0.35 Vocab. 0.39 Comp.	Grade level,* Implementer, Integration of reading/writing, Outcome measure type,* Literacy difficulty, Number of instructional sessions, Proportion of instruction devoted to reading vs. writing,* Professional development, Published vs. non-published,* Sample size

*Note.* E = experimental. QED = Quasi-experimental design. SWRD = students with reading disabilities. SWLD = students with learning disabilities; LD = learning disability; S = students; Para/Vol = paraprofessional/volunteer. PA= phonological awareness; Comp = comprehension; Word recog = word recognition; Vocab = vocabulary; Found = foundational (includes PA and word reading); CF = code focused (PA, phonics and fluency); NA= not applicable; NR= not reported.

\* = statistically significant

<sup>a</sup> = statistically significant for foundational outcomes 1980-2004

<sup>b</sup> = statistically significant for comprehension outcomes 2005-2011

<sup>c</sup> = statistically significant for foundational outcomes

<sup>d</sup> = statistically significant for comprehension outcomes