Hammill Institute on Disabilities

Remedial and Special Education
2019, Vol. 40(3) 131–149

DOI: 10.1177/0741932517749940
rase.sagepub.com
SAGE

A Review of Summarizing and Main Idea Interventions for Struggling Readers in Grades 3 Through 12: 1978–2016

Elizabeth A. Stevens, MA Ed[1], Sunyoung Park, MA[1], and Sharon Vaughn, PhD[1]

Abstract

This systematic review examines the effects of summarizing and main idea interventions on the reading comprehension outcomes of struggling readers in Grades 3 through 12. A comprehensive search identified 30 studies published in peer-reviewed journals between 1978 and 2016. Studies included struggling reader participants in Grades 3 through 12; targeted summarizing or main idea instruction; used an experimental, quasi-experimental, or single-case design; and included a reading comprehension outcome. A meta-analysis of 23 group design studies resulted in a statistically significant mean effect of 0.97. Group size, number of sessions, grade level, and publication year did not moderate treatment effect. Visual analysis of six single-case designs yielded *strong evidence* for retell measures and a range of evidence for short-answer comprehension measures. Findings suggest that main idea and summarizing instruction may improve struggling readers' main idea identification and reading comprehension. Limitations include the lack of standardized measures and the unreported, changing description of the counterfactual.

Keywords

reading comprehension, intervention, main idea, summarizing, struggling reader, learning disability, middle school, high school

As students progress through middle and high school, they must be able to read and understand content presented through text (Hagaman, Casey, & Reid, 2016). Proficient readers integrate several processes, such as fluent decoding and purposeful strategy use, to monitor their understanding while reading (Hagaman et al., 2016). Reading comprehension is a critical life skill because it enables students to be successful in society, the workforce, and social situations (Biancarosa & Snow, 2006). Students need advanced literacy skills to compete in a global economy, as well as the flexibility to apply those skills in a variety of contexts (e.g., technology; Biancarosa & Snow, 2006).

In spite of the importance of reading comprehension, many students lack the reading skills necessary to be successful in postsecondary education and employment (Kamil et al., 2008). Only 36% of fourth graders, 34% of eighth graders, and 37% of 12th graders read at a proficient level (National Center for Education Statistics [NCES], 2015), suggesting that less than half of students have acquired the necessary skills to read and understand grade-level texts. Reading performance is particularly low for students with identified disabilities. In 2015, 12% of fourth graders, 8% of eighth graders, and 12% of 12th graders with disabilities performed at a proficient level in reading (NCES, 2015). The National Longitudinal Transition Study 2, which examined 10 years of high school transcript data for secondary students with disabilities, confirmed that students with disabilities lag behind general education peers in academic performance (Newman et al., 2011). Many students are fluent word readers but simply do not comprehend text (Biancarosa & Snow, 2006). This suggests a critical need for targeted comprehension intervention for upper elementary, middle, and high school students. Despite this need, many teachers view themselves as content area teachers rather than reading teachers and assume that students with adequate word reading skills also have adequate reading comprehension (Edmonds et al., 2009). Consequently, formal reading instruction is often lacking beyond the elementary years (Kamil et al., 2008).

[1]The University of Texas at Austin, USA

Corresponding Author:
Elizabeth A. Stevens, College of Education SZB 228, The University of Texas at Austin, 1912 Speedway D4900, Austin, TX 78712-1284, USA.
Email: elizabeth.stevens@utexas.edu

Experts have identified main idea generation and summarization as effective instructional practices for improving students' literacy (Biancarosa & Snow, 2006; Goldman, 2012; Kamil et al., 2008). Identifying main ideas is an active, meaning-making process that facilitates comprehension because it helps the reader remember important information and develop a global understanding of the text (Hagaman et al., 2016; Jitendra, Chard, Hoppes, Renouf, & Gardill, 2001; Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007). High-stakes assessments based on rigorous state and national standards (e.g., Common Core State Standards; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) require students to identify paragraph-level main ideas, explain how those main ideas are supported by key details, and summarize narrative and expository texts.

Summarization Processing Models

Microprocessing

Kintsch and van Dijk (1978) proposed a summarization model in which the reader develops a macrostructure, or global understanding of the text, using the text's microstructure (i.e., meaning units that comprise the text's base). The reader applies specific subskills to the text's microstructure to form a gist (i.e., the overarching main idea of the text, which can be expressed in one to two sentences). First, the reader deletes trivial and redundant information, distinguishing important information from unnecessary details. Next, the reader generalizes information, using a superordinate, categorical name or action for a list of items (e.g., *farm animals* for cows, pigs, horses, and chickens) or subcomponent actions (e.g., *Mary went on a trip* for Mary packed a bag, Mary got into a taxi, Mary went to the airport, etc.). Finally, the reader undergoes the process of construction, or selecting a main idea sentence explicitly stated in the text; if one is not available, the reader invents the implied main idea sentence. In other words, this model involves a ground-up process of identifying *and* condensing important meaning units to arrive at the main idea.

Macroprocessing

In contrast, Meyer, Brandt, and Bluth (1980) suggested that readers arrive at the gist using macrostructure cues or the text structure organization. In this top-down approach, specific text features (e.g., headings or a topic sentence that explicitly states the main idea) facilitate the reader's identification of the text's macrostructure. The reader approaches the text with an existing pattern in mind (e.g., problem/solution, compare/contrast). During reading, the reader integrates information into that preexisting structure to extract the gist (Meyer et al., 1980).

Examining Micro- and Macro-Based Strategy Use

Brown and Day (1983) investigated students' application of Kintsch and van Dijk's (1978) rules to develop the macrostructure of a text. Students as young as fifth grade were able to delete trivial or redundant information but were less likely to accurately categorize information (i.e., generalization) or to identify implicit main ideas in the text (i.e., construction). Students in seventh grade, 10th grade, and college also struggled to invent a topic sentence. In contrast, Meyer and colleagues (1980) investigated macro-based instruction, finding that ninth graders with good comprehension skills utilized text structure more than students with poor comprehension skills. This top-down approach improved students' discrimination between important and irrelevant information and overall recall of important information. Similarly, Gallini, Spires, Terry, and Gleaton (1993) examined micro- and macro-based strategy use for struggling high school readers. Students received schema-based instruction focused on the top-level structure of the passage (i.e., diagrammatic representations of the hierarchical relationships among concepts and ideas across paragraphs in a given passage), micro-based instruction focused on making connections within and across sentences, or traditional reading instruction. The macro-based condition outperformed the micro-based and control conditions on immediate and delayed measures of reading comprehension.

Students' use of micro- or macro-based strategies may differ as a function of text type. Expository text structure (e.g., problem/solution or compare/contrast) and text features (e.g., headings) enable the use of macro-based strategies to identify main ideas. However, students may employ micro-based strategies for narrative texts because the structure lacks the same organizational features found in expository texts. If macrocues are unavailable, the reader may resort to micro-based processing of the text. This suggests the need for explicit instruction in micro- and macro-based strategies, so that struggling readers can flexibly apply both processes (Gallini et al., 1993).

Identifying Main Ideas and Summarizing Present Challenges for Struggling Readers

The complexity and challenge of identifying main ideas cannot be overestimated. Summarizing is a difficult skill because it requires readers to actively monitor their understanding and simultaneously identify important information, eliminate irrelevant details, and integrate main ideas across paragraphs and chapters (Duke & Pearson, 2008; Jitendra et al., 2001; Watson, Gable, Gear, & Hughes, 2012). Monitoring for meaning (e.g., stopping while reading to reflect on the paragraph- or passage-level main ideas) is

particularly difficult for students with reading disabilities, as they may not realize when meaning breaks down and take the necessary steps to repair misunderstandings (Miller, Darch, Flores, Shippen, & Hinton, 2011). Brown and Day (1983) found that upper elementary, middle school, high school, and even college students struggled to identify main ideas in text. In particular, students with learning disabilities have difficulty identifying and including main ideas in summaries; instead, students may include unfamiliar or unusual information they find particularly interesting (Williams, 2006). Generalizing from the ideas in the text to write a main idea topic sentence may also be a challenge, even when the title of the passage clearly provides the main idea (e.g., listing the ways birds keep warm in winter without specifying "this is about how birds keep warm in winter"; Taylor, 1986).

Recent Syntheses and Meta-Analyses

Prior reviews examined the effects of reading interventions for students with reading difficulty in upper elementary, middle, and secondary settings. For example, Scammacca, Roberts, Vaughn, and Stuebing (2015) reported the effects by type of intervention (i.e., word study, vocabulary, reading comprehension, fluency, multiple components), revealing a large effect of 0.74 for reading comprehension interventions and a small effect of 0.20 for multicomponent interventions. The authors did not examine specific components of reading comprehension strategy instruction (e.g., main idea or summarization instruction) within reading comprehension interventions. Solis et al. (2012) synthesized reading comprehension interventions, revealing large effects on researcher-developed measures for main idea and summarization strategy instruction; however, this review included only reading comprehension interventions provided to students with learning disabilities in Grades 6 through 8. Although current evidence supports reading intervention for students with reading difficulty or learning disabilities, studies targeting solely main idea or summarization instruction for struggling readers across Grades 3 through 12 have not been examined. Due to the importance of summarization and main idea instruction for reading comprehension, this review aims to identify effective practices for teaching these strategies to students with reading difficulties.

Purpose and Research Questions

Commercial reading programs (e.g., Scott Foresman; Jitendra et al., 2001) and evidence-based, multicomponent reading interventions (e.g., Collaborative Strategic Reading; Vaughn et al., 2011) often include main idea instruction, yet prior reviews have not examined the effects of such instruction (i.e., as the sole component of the intervention vs. as part of a multicomponent intervention) on struggling readers' comprehension outcomes. Furthermore, instruction often consists of telling students to find the most important idea without explicitly teaching students *how* to do so (Williams, 2006). Understanding the effects of summarizing and main idea interventions is needed to improve the quality of reading instruction for struggling readers, which may result in improved reading comprehension and potentially enhance students' opportunity for success later in life. In addition, improved understanding of the effects of summarizing and main idea interventions may provide insight on ways to strengthen the reading comprehension component of multicomponent interventions and elucidate future research needs. This systematic review addressed the following primary research question:

> **Research Question 1:** What are the effects of summarizing and main idea interventions on the reading comprehension outcomes of struggling readers in Grades 3 through 12?

In addition, we aimed to answer the following secondary research questions:

> **Research Question 2:** What instructional practices exist in the literature to address summarizing and main idea? What type of text, narrative, or expository is used in summarization and main idea instruction?

Method

Operational Definitions

Summarizing refers to a skill in which students identify the most important elements of a passage (i.e., the main ideas) in their own words. Students must be able to distinguish between the main ideas and details provided in the passage. Finally, the main ideas are combined to provide a brief synopsis, written or spoken, of the most important aspects of the text. *Main idea* refers to the most important idea within a particular section, or paragraph, of the passage. Identifying the main idea is a critical subskill in the summarization process. A *struggling reader* refers to a student identified with a learning disability, a reading disability or difficulty, or at risk for reading difficulty.

Search Procedures

We conducted a computer search of three electronic databases: Education Source, Educational Resources Information Clearinghouse (ERIC), and PsycINFO. The search was limited to peer-reviewed journals published between January 1978 and March 2016. We began the search in 1978 after the release of Durkin's observation study, examining the state of reading comprehension

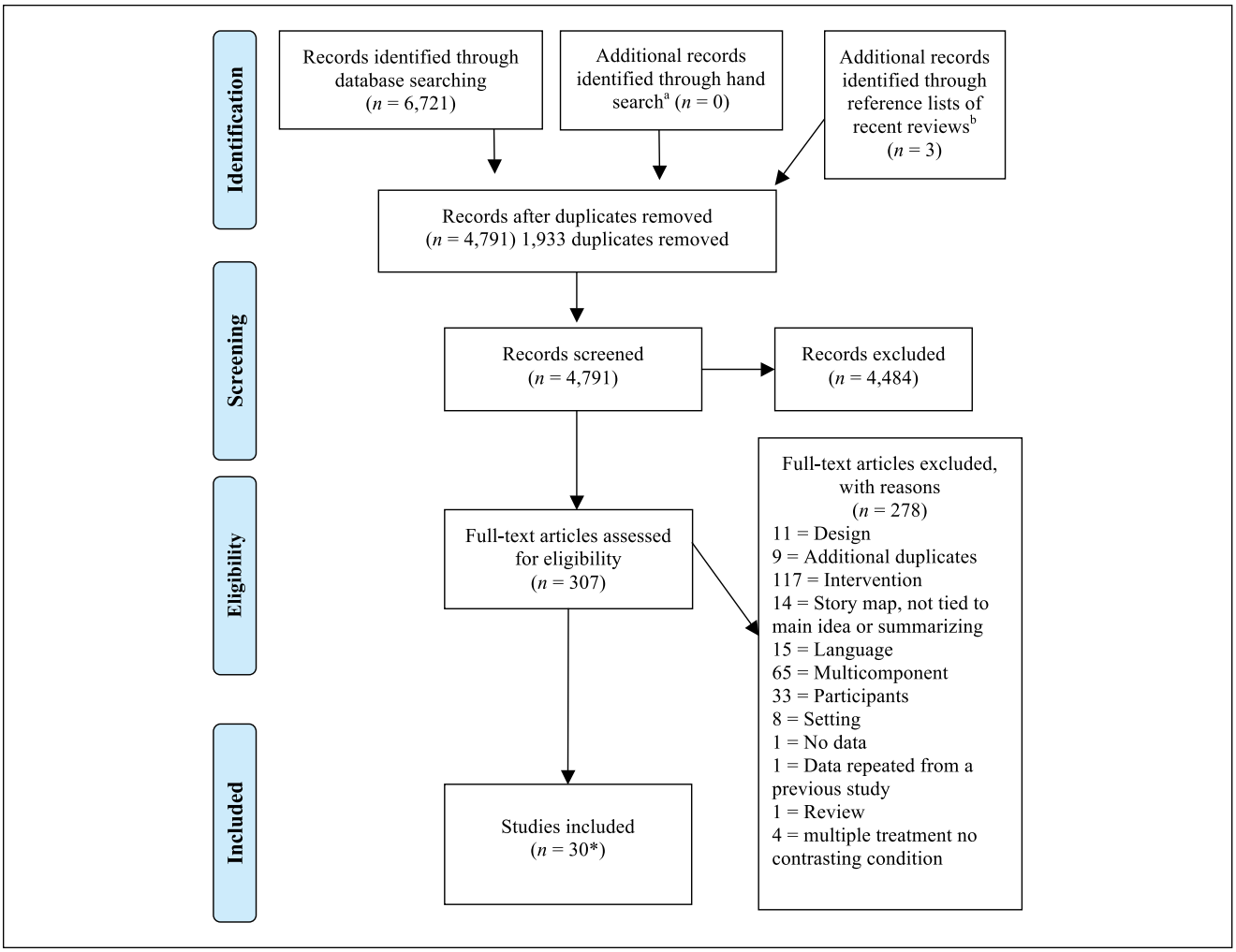

Figure 1. PRISMA diagram detailing the search process.
[a]Exceptional Children, Journal of Educational Research, Journal of Learning Disabilities, The Journal of Special Education, Learning Disabilities Quarterly, Learning Disabilities Research and Practice, Scientific Studies of Reading, Reading Research Quarterly, and Remedial and Special Education.
[b]Mason (2013); Scammacca et al. (2015); Solis et al. (2012).
[c]Weisberg and Balajthy (1990) and Schunk and Rice (1992) contributed two studies each.

instruction in third through sixth grade reading and social studies classrooms. Findings from the observation study revealed a dearth (i.e., less than 1% of the instructional time observed) of explicit comprehension instruction (Durkin, 1978). This seminal work provided the impetus for a lasting line of research targeting reading comprehension instruction (e.g., Brown & Day, 1983; Jitendra, Hoppes, & Xin, 2000). We used the following search terms: *read**, and *comprehen** OR *"main idea*"* OR *summariz** OR *paraphras** OR *"paragraph shrinking,"* and *disabilit** OR *disorder* OR "*struggling reader**" OR *"learning problem*"* OR *"at risk"* OR *"high risk"* OR *difficult** OR *delay** OR *"poor read*,"* and *interven** OR *teach** OR *instruct** OR *strateg** OR *program** OR *train**. See Figure 1 for a PRISMA diagram detailing the search process (i.e., Preferred Reporting Items for Systematic Reviews and Meta-Analyses; Moher, Liberati, Tetzlaff, & Altman, 2009). Thirty studies (i.e., 28 publications) met the inclusion criteria.

Inclusion Criteria

We included studies that met the following criteria:

1. Published in English in a peer-reviewed journal from January 1978 to March 2016.
2. Employed an experimental, quasi-experimental, multiple treatment, or single-case design (SCD) providing a treatment and comparison to determine experimental effect (i.e., single-group pretest/post-test, AB single-case, qualitative, and case study designs were excluded). Multiple treatment studies were included if one of the treatments served as a

contrast to the treatment of interest (i.e., summarizing or main idea instruction).

3. Included participants identified as struggling readers in Grades 3 through 12. Struggling readers were defined as (a) students identified with learning disabilities; (b) students identified with reading disabilities; or (c) students with reading difficulty or at-risk status as determined by low performance on a reading measure (e.g., Comprehensive Test of Basic Skills, Illinois State Achievement Test, Degree of Reading Progress), placement in a remedial reading class, or district or school identification (e.g., the principal and teachers selected students struggling with reading comprehension). Studies with additional participants (i.e., students in kindergarten through Grade 2 or students without reading difficulty) were included if at least 50% of the sample included the targeted population (i.e., struggling readers in Grades 3 through 12), or disaggregated data were provided for these students. We included English language learners, students with behavioral disorders, and students with attention deficit/hyperactivity disorder if they were also identified as struggling readers. We excluded studies targeting students with autism, intellectual disabilities, and vision or hearing impairments. We excluded Grades K through 2 as targeted reading instruction for students with reading difficulties at these grade levels typically emphasizes the foundation skills of reading (i.e., phonemic awareness, phonics, and fluent word reading). As students become proficient word readers, the focus of reading instruction shifts from "learning to read" to "reading to learn" (Chall, 1983). For this reason, we restricted our search to Grades 3 through 5 at the elementary level and all grades at the middle and high school levels.
4. Examined a reading intervention, provided in English, targeting summarizing or main idea instruction. Story map interventions were included if used as a mechanism for summarization or retelling. We excluded multicomponent interventions (i.e., interventions targeting reading comprehension, word reading or decoding, reading fluency, and/or vocabulary) and summarizing or main idea interventions implemented in a listening comprehension format.
5. Provided instruction as part of the school programming (i.e., home, clinic, and camp settings were excluded).
6. Included at least one dependent variable assessing reading comprehension outcomes.

Coding Procedures

Studies that met the inclusion criteria were coded using a protocol (Vaughn, Elbaum, Wanzek, Scammacca, & Walker, 2014) developed for education-related intervention research based on study features described in the What Works Clearinghouse (WWC) Design and Implementation Assessment Device (Valentine & Cooper, 2008), and used in previous meta-analyses (e.g., Wanzek et al., 2013). We extracted the following data from the group and SCD studies: (a) participant information (e.g., age, grade level, number of participants, type of struggling readers), (b) research design, (c) treatment fidelity (i.e., coding whether the authors reported a fidelity of treatment check), (d) description of treatment and comparison group(s) or baseline and intervention phase(s), (e) total sessions and hours of intervention provided, and (f) measures. We also coded for the clarity of causal inference (e.g., differential attrition between intervention and comparison groups, adequate equating procedures for quasi-experimental designs, any indication of the plausibility of intervention contaminants) and results and effect sizes (ESs) in the group design studies. We extracted the following additional data from the SCD studies using the WWC standards for design and evidence evaluation: (a) systematic manipulation of the independent variable, (b) interobserver agreement, (c) the number of attempts to demonstrate an intervention effect, (d) the number of data points per phase, and (e) data extraction and visual analysis from the graphs provided (Kratochwill et al., 2013; see the SCD analysis section for more details). We used the gold standard method (Gwet, 2001) to establish interrater reliability prior to coding. The first author, a researcher with experience using and publishing syntheses with the codesheet, provided an initial 4.5-hr training session to two graduate research assistants studying reading intervention research. In addition, the first author and one graduate research assistant had prior graduate-level training in conducting visual analysis of SCD studies. The researcher described the codesheet and modeled each step of the coding process for a sample intervention study, and then the research assistants practiced coding additional intervention studies of different design types. Upon completion of the training, the coders independently coded a study to establish reliability. Coders achieved an interrater reliability score of .96, determined by the number of items in agreement divided by the total number of items. After establishing initial reliability, each study was independently coded by two coders. The coders met to review each codesheet, and to identify and resolve any discrepancies. When the coders were unable to resolve a specific code, the first author reviewed the study, and the author team made final decisions by consensus.

Meta-Analysis Procedures for the Group Design Studies

For experimental, quasi-experimental, and multiple treatment designs, ESs were calculated as the difference between the groups' means divided by the pooled standard deviation; Hedge's g is reported to provide a less biased

estimate of ES with particularly small samples. We used R studio with "robumeta" package program for each treatment and comparison contrast on all comprehension outcomes (Fisher, Tipton, & Zhipeng, 2017). To address dependency in the data resulting from multiple outcome measures or multiple treatment groups per study, we used robust variance estimation (RVE; Hedges, Tipton, & Johnson, 2010). RVE uses a mean correlation value (i.e., the correlations between the ESs in each study are usually unreported), ρ, to calculate study weights and between-study variance. Rather than averaging the ES estimates per study or selecting one ES per study, this technique uses all available outcome data and provides more precise standard errors (Hedges et al., 2010). We conducted a sensitivity analysis with various ρ values and found that the results did not differ based on the selected value; the results are based on a ρ value of .80. We also used a small-sample correction, adjusting each coefficient's degrees of freedom to address inflated Type I error rates for meta-analyses with less than 40 studies (Tipton, 2013).

The heterogeneity of variance in ESs was evaluated using I^2 and τ^2 statistic. In the presence of statistically significant variability between ESs, we conducted four additional, separate meta-regression models with the moderator as the covariate. We were unable to conduct one meta-regression model with the four covariates of interest due to the small number of studies included in the meta-analysis, and not all studies reported information for each moderator variable (e.g., one study did not report number of sessions). Moderators included three categorical variables and one continuous variable: total sessions (12 or fewer sessions vs. 13 sessions or more), group size (one to four vs. five or more), grade (elementary vs. middle and high school), and publication year (i.e., centered to interpret the intercept as the pooled ES given the average year of publication). Most studies reported grade, number of sessions, or group size as a range; as such, categorical analyses were conducted to maximize the number of ESs included in each analysis (i.e., we used the median value to determine the cutoff for number of sessions and group size). Studies that reported a range of group size or sessions and could be placed in either category were excluded from the moderator analysis (e.g., Gajria & Salvia, 1992, reported a range of 10–19 sessions). In each meta-regression model, we dummy coded the first level in the category as 0 (i.e., 12 or fewer sessions, group size of one to four, and elementary grades) and the second level in the category as 1 (i.e., 13 or more sessions, group size of five or more, middle and high school grades). To avoid an increase in Type I error as a result of running four RVE regression models, we used the Bonferonni correction, adjusting the *p* value for statistical significance to .0125 for each analysis (i.e., .05 divided by 4; Abdi, 2007).

Analysis Procedures for the Synthesis of SCD Studies

We applied the WWC two-step process for evaluating SCD research: We (a) evaluated the study as *meets design standards, meets design standards with reservations*, or *does not meet design standards* (see Table 3 for evaluation results); for studies rated as *meets design standards* or *meets design standards with reservations*, then we (b) evaluated the level of evidence (Kratochwill et al., 2013). A study received a rating of *meets design standards* if the following criteria were met: (a) the independent variable was systematically manipulated; (b) each outcome variable was systematically measured over time by more than one assessor, with interobserver agreement exceeding 0.80 on at least 20% of the data points; (c) experimental control was demonstrated if the design provides at least three different opportunities to demonstrate an intervention effect at different time points (i.e., at least three baseline and three intervention phases in a multiple baseline design); and (d) the phase included a minimum of five data points. If a multiple baseline design met the aforementioned criteria and included at least three to four data points per phase, then the study received a rating of *meets design standards with reservations*. A study that did not meet criteria a, b, or c, or contained fewer than three data points per phase was rated *does not meets design standards*. WWC recommends conducting visual analysis to evaluate the effects within SCDs for studies that *meet design standards* or *meet design standards with reservations* (Kratochwill et al., 2013). The following steps were applied to visually examine the within- and between-phase data for each study: We (a) examined baseline phase data to determine if a predictable and stable pattern exists; (b) examined within-phase patterns, including the level, trend, and variability of the data; (c) compared the data in each phase with data in adjacent phases to assess whether the introduction of the independent variable was associated with a predicted change in the dependent variable (i.e., immediacy of the effect, overlap, consistency of data patterns across similar phases); and (d) integrated the information from all phases, and identified whether there were at least three demonstrations of an effect at three different time points (Kratochwill et al., 2013). A study received a *strong evidence* rating if it provided at least three demonstrations of an effect at different time points; if a study did not provide at least three demonstrations of an effect, then it received a *no evidence* rating. Finally, a study received a *moderate evidence* rating if at least three demonstrations of an effect were present with at least one demonstration of a noneffect. ES estimation followed for those studies that received *strong evidence* and *moderate evidence* ratings.

ES estimates for SCD studies. We used two nonoverlap indices to evaluate SCD studies with *strong* or *moderate* evidence (Kratochwill et al., 2010). We selected the percentage of nonoverlapping data (PND) because it remains a commonly used metric in the field (Parker, Vannest, & Davis, 2014). The total number of data points during the intervention phase that exceeded the highest baseline data point divided by the total number of data points within that phase and multiplied by 100 provided the PND for each comprehension outcome (i.e., data from the postinstruction phase were used in four studies as data during the treatment were not available; Hagaman & Reid, 2008; Hagaman, Casey, & Reid, 2012; Hagaman et al., 2016; Mason, Snyder, Sukhram, & Kedem, 2006; Scruggs, Mastropieri, & Casto, 1987). PND results were interpreted as follows: 90% or greater is highly effective, 70% to 90% is moderately effective, 50% to 70% is minimally effective, and 50% or less is ineffective (Scruggs & Mastropieri, 1998). One potential limitation of PND is that it is calculated using the most extreme baseline data point. As such, we provided an additional ES metric, nonoverlap of all pairs (NAP), which reports the probability that a data point randomly selected from the treatment phase will exceed a point randomly selected from the baseline phase (Parker & Vannest, 2009). Each baseline phase data point was compared with each treatment phase data point; an overlapping pair was defined as a pair of data points in which the baseline data point was higher than the treatment data point. Overlapping pairs were counted as 1 point, whereas a tied pair was counted as half a point. NAP was reported as a percentage, calculated by subtracting the number of overlapping pairs from the total possible pairs (i.e., the number of data points in baseline multiplied by the number of points in intervention) and dividing the total possible pairs. NAP differs from PND, in that it evaluates all possible data overlap between the two phases (Parker & Vannest, 2009). We applied the following guidelines for interpreting NAP: 0% to 65% is considered a weak effect, 66% to 92% is considered a medium effect, and 93% to 100% is considered a strong effect.

After conducting the design and evidence evaluation for each study, we used the software program WebPlotDigitizer—a recommended program based on a recent examination of the reliability of numeric coding, the validity of the data extraction compared with real data, and the overall usability of the program—to extract data from the graphs (Moeyaert, Maggin, & Verkuilen, 2016; Rohatgi, 2015). A screenshot of each graph was opened in WebPlotDigitizer, the x and y axes were calibrated, the coder clicked on each data point in the graph, and the coordinates of each data point were exported to Microsoft Excel 2011. Next, the y values for the baseline and intervention phases were pasted into the single-case ES calculator, an online tool we used to compute the two nonoverlap indices for each case (Pustejovsky, 2017). PND and NAP values are reported in Table 2S (see Supplemental Material).

Results

The results are reported in three sections: (a) findings from the meta-analysis and moderator analyses of the group design studies, (b) the instructional practices addressed in the group design studies, and (c) a synthesis of the SCD results and instructional practices. Thirty studies met inclusion criteria for this review: 24 group design studies (i.e., 20 treatment–control experiments, three multiple treatment experiments, one quasi-experiment) and six SCD studies. Table 1 presents descriptive characteristics (e.g., publication year, grade level, group size, interventionist) of the 30 group and SCD studies. This corpus represents a total of 983 participants with a range of two to 81 participants per study. Table 2 describes the treatment and comparison conditions for the group design studies, and Table 1S (see Supplemental Material) presents the results of those studies. Table 3 describes the phases and participants for the SCD studies, and Table 2S (see Supplemental Material) reports the PND and NAP for those studies.

Meta-Analytic Results of the Group Design Studies

A total of 109 posttest ESs were reported from 23 group design studies (i.e., we were unable to calculate ESs for one study; Boyle & Weishaar, 1997). The RVE random-effects model estimated a statistically significant, large treatment effect of 1.25 ($p < .001$, $SE = 0.22$). The ESs ranged from −0.99 to 5.98. Due to the wide range of estimates and standard errors, we conducted a sensitivity analysis based on Lipsey and Wilson's (2001) definition of extreme values (i.e., exceeding 3 standard deviations above and below the mean value). After excluding six outlier ESs that exceeded 3.0 standard deviations above the mean, the remaining 103 ESs resulted in a statistically significant, large effect of 0.97 ($p < .001$, $SE = 0.14$; 95% confidence interval [CI] = [0.68, 1.25]). We consider this a more precise estimate of the mean effect.

We detected significant variability between ESs ($I^2 = 72.34$, $\tau^2 = .37$), so we conducted four separate, exploratory regression models with each moderator variable of interest (see Table 4). Three of the moderator variables, number of sessions, group size, and grade level, were not statistically significant predictors of ES; although the difference between the categories was not significant, summarizing and main idea interventions resulted in significant ES estimates for each category. Finally, publication year did not significantly predict ES ($\beta = -.001$, $p > .5$).

Table 1. Descriptive Characteristics for Experimental (n = 24) and Single-Case (n = 6) Design Studies.

	Experimental		Single-case	
Characteristic	*n*	%	*n*	%
Publication year				
1980s	5	21		
1990s	13	54	2	33
2000s	3	13	1	17
2010–present	3	13	3	50
Measure type				
Unstandardized	96	93	12	100
Standardized	7	7		
Grade level				
Elementary (3–5)	7	29	1	17
Middle (6–8)	6	25	3	50
High (9–12)	6	25	2	33
Elementary and middle (3–8)	2	8		
Middle and high (7–12)	3	13		
M Group size[a]				
One-on-one	3	13	4	67
2–4	3	13	2	33
5 or more	5	21		
NR	11	46		
Interventionist				
Teacher	5	21		
Researcher	14	58	6	100
NR	5	21		

Note. NR = not reported.
[a]Dimino et al. (1990) and Gallini et al. (1993) not included because they reported group sizes in two categories.

Publication bias. We applied the Trim and Fill method to evaluate the presence of publication bias in our results (Duval & Tweedie, 2000). This method removes the estimates causing asymmetry in the funnel plot, calculates a mean effect after removing those estimates, then replaces the removed studies along with additional ESs to correct the asymmetry. The method estimates the number of studies potentially missing from the plot and provides an adjusted mean effect that includes the missing values. Results of the Trim and Fill analysis indicated that publication bias did not affect the mean estimate.

Summarizing and main idea instructional practices and text type. We aimed to identify the instructional practices and text type used in summarizing and main idea research.

Instructional practices. Four studies examined the use of text structure instruction on students' main idea identification or summarization (Bakken, Mastropieri, & Scruggs, 1997; Dimino, Gersten, Carnine, & Blake, 1990; Miller et al., 2011; Weisberg & Balajthy, 1989). Students learned to distinguish among expository organizational structures, recognize text structure signal words, or attend to structural cues (e.g., topic and concluding sentence) to identify main ideas and summarize text. Six studies examined the effects of paraphrasing strategies on main idea identification (Bakken et al., 1997; Ellis & Graves, 1990; Jenkins, Heliotis, Stein, & Haynes, 1987; Katims & Harris, 1997; Mason, 2004; Mastropieri, Scruggs, Spencer, & Fontana, 2003). Students were taught a three-step process (e.g., read a paragraph, ask yourself what was the main idea and two details, put the main idea and details into your own words) or a two-step process (e.g., identify the most important "who" or "what" in the text, identify the most important thing about the "who" or "what," and write the summary sentence) for paraphrasing main ideas. Three studies implemented graphic organizers (i.e., cognitive mapping strategies or story structure diagrams) to facilitate main idea identification and summary writing (Boyle, 1996; Boyle & Weishaar, 1997; Faggella-Luby & Wardwell, 2011).

Seven studies investigated microprocessing strategies on text summarization (Gajria & Salvia, 1992; Gallini et al., 1993; Mason, 2004; Schunk & Rice, 1992; Weisberg & Balajthy, 1990). Three of these studies examined the effects of instruction on variations of Brown and Day's (1983) five summarization rules: (a) delete trivial information; (b) delete redundant information; (c) generalize information using a superordinate, categorical name; (d) select the main idea topic sentence from the text; and (e) invent the main idea sentence if one is not explicitly stated (Gajria & Salvia, 1992; Mason, 2004; Weisberg & Balajthy, 1990). In Gallini et al. (1993), the micro-based condition provided explicit instruction on anaphoric relations (i.e., the relationship between an anaphor and its antecedent) and connectives (e.g., a word that connects phrases or sentences) to link ideas across sentences. In two experiments, Schunk and Rice (1992) used a five-step comprehension strategy to encourage students to make connections among ideas across sentences: (a) read the questions, (b) read the passage to find out what it is mostly about, (c) think about what the details have in common, (d) think about what would make a good title, (e) reread the story if you do not know the answer to a question.

Finally, five studies evaluated the effects of main idea or summarizing instruction combined with self-monitoring (Graves, 1986; Jitendra et al., 2000; Malone & Mastropieri, 1992; Mason, 2013; Wong & Jones, 1982). Self-monitoring strategies included goal setting during summary writing (e.g., pick goals for an essay, list ways to meet goals, make notes, and sequence the notes; Mason, 2013), self-questioning techniques (e.g., Do I understand what the whole story is about? Graves, 1986), and cue cards to check off completed steps in the strategy (Jitendra et al., 2000; Malone & Mastropieri, 1992).

Text type. Eleven studies did not report text type used in the intervention. Ten used expository text, four used narrative text, and five used expository and narrative texts.

Table 2. Experimental Study Information.

Study	Design	*N*	Grade	Struggling reader	Text type	Description of conditions	Total sessions	Total hours	Fidelity reported
Bakken, Mastropieri, & Scruggs (1997)	Treatment comparison experiment	54	8	LD	T1: E T2: N, E	T1: Text structure-based strategy T2: Paragraph restatement strategy C: Traditional instruction	3	1.57	N
Borkowski, Weyhing, & Carr (1988)	Multiple treatment experiment	73	Upper elem	RD, LD	NR	T1: Summarization with complex attribution instruction T2: Summarization with attribution instruction T3: Summarization instruction T4/C: No explicit strategy or attribution instruction	5	NR	N
Boyle (1996)	Treatment comparison experiment	30	6–8	LD	NR	T1: TRAVEL cognitive mapping strategy C: Typical reading instruction	6	5	N
Boyle & Weishaar (1997)	Treatment comparison experiment	39	10–12	LD	NR	T1: TRAVEL student-generated cognitive mapping T2: Expert-generated cognitive mapping C: Blank paper for note taking while reading	10	8.3	N
Dimino, Gersten, Carnine, & Blake (1990)	Treatment comparison experiment	32	9	R-dif, LD	N	T1: Story grammar instruction to facilitate written retell C: Traditional literature instruction	19	NR	Y
Ellis & Graves (1990)	Treatment comparison experiment	32	5–7	LD	N, E	T1: Repeated reading (RR) T2: Paraphrasing strategy instruction T3: Paraphrasing strategy instruction with RR C: Reading instruction without main idea instruction or RR	8	1.6	N
Faggella-Luby & Wardwell (2011)	Treatment comparison experiment	81	5 and 6	A-R	N	T1: Story structure routine T2: Sustained silent reading C: Typical practice Tier 2 intervention	35–48	17.5–24	Y
Gajria & Salvia (1992)	Treatment comparison experiment	30	6–9	LD	E	T1: Summarization instruction C: NR	10–19	6.5–11	N
Gallini, Spires, Terry, & Gleaton (1993)	Treatment comparison experiment	66	High school	R-dif	E	T1: Explicit macroprocessing strategy instruction T2: Explicit microprocessing strategy instruction C: Status quo remedial reading program	12	10	N
Graves (1986)	Treatment comparison experiment	24	5–8	LD	NR	T1: Direct instruction on finding main ideas with self-monitoring T2: Direct instruction on finding main ideas C: No direct instruction on finding main ideas	NR	NR	N

(continued)

Table 2. (continued)

Study	Design	*N*	Grade	Struggling reader	Text type	Description of conditions	Total sessions	Total hours	Fidelity reported
Jenkins, Heliotis, Stein, & Haynes (1987)	Treatment comparison experiment	32	3–6	LD	N	T1: Restatement instruction C: Regularly assigned seat work activities	10–15	3.3–5	N
Jitendra, Hoppes, & Xin (2000)	Treatment comparison experiment	33	6–8	LD	NR	T1: Main idea strategy instruction on rules for identifying details, the main person or group, and the central action in a paragraph; self-monitoring card to check strategy use C: Typical special education resource instruction	15	7.5–10	Y
Katims & Harris (1997)	Treatment comparison experiment	25	7	LD	N, E	T1: Paraphrasing strategy C: District-mandated Reading Workshop	13	4.3	N
Malone & Mastropieri (1992)	Treatment comparison experiment	45	6–8	LD	NR	T1: Summarizing instruction T2: Summarizing instruction with self-monitoring C: Traditional reading comprehension instruction	2	NR	N
Mason (2004)	Multiple treatment experiment	32	5	R-dif, LD	E	T1: TWA; RAP strategy used after reading to identify main ideas T2/C: Reciprocal questioning	11–15	3.7–5	Y
Mason (2013)	Treatment comparison experiment	77	4	R-dif, LD	E	T1: TWA; RAP strategy used after reading to identify main ideas T2: TWA + PLANS to develop and evaluate goals related to summary writing C: Opportunity to read; no treatment	18–22	9–11	Y
Mastropieri, Scruggs, Spencer, & Fontana (2003)	Multiple treatment experiment	16	10 and 12	LD	E	T1: Peer tutoring with partner reading and summarization strategies T2/C: Guided notes	18	27	Y
Miller, Darch, Flores, Shippen, & Hinton (2011)	Treatment comparison experiment	38	3–5 (*N* = 32); 6–7 (*N* = 6)	LD	NR	T1: Explicit instruction C: Basal instruction	12	9	Y
Schunk & Rice (1992) Experiment 1	Treatment comparison experiment	33	4 and 5	R-dif	N, E	T1: Strategy instruction to identify main ideas T2: Strategy instruction with specific strategy-use feedback C: Reading instruction without explicit comprehension strategy instruction	15	8.75	N

(continued)

Table 2. (continued)

Study	Design	*N*	Grade	Struggling reader	Text type	Description of conditions	Total sessions	Total hours	Fidelity reported
Schunk & Rice (1992) Experiment 2	Treatment comparison experiment	33	4 and 5	R-dif	NR	T1: Strategy instruction to identify main ideas T2: Strategy instruction to identify main ideas and details C: Reading instruction without explicit comprehension strategy instruction	20	NR	N
Weisberg & Balajthy (1989)	Treatment comparison quasi-experiment	32	10–12	R-dif	E	T1: Instruction on text structure (e.g., compare/contrast), construct graphic organizer, synthesize information into a written summary C: Reading and teacher-led discussion	6	4	N
Weisberg & Balajthy (1990) Experiment 1	Treatment comparison experiment	24	High school	RD	E	T1: Summarizing instruction and practice applying the strategy T2: Summarizing instruction only C: Regular reading instruction	3	NR	N
Weisberg & Balajthy (1990) Experiment 2	Treatment comparison experiment	12	High school	RD	E	T1: Summarizing instruction and practice applying the strategy C: Regular reading instruction	4	NR	N
Wong & Jones (1982)	Treatment comparison experiment	60	8 and 9	LD	NR	T1: Self-questioning instruction to identify passage main ideas C: Students read passages	2	4	N

Note. *N* = sample size; LD = learning disability; T1 = Treatment 1; E = expository text; T2 = Treatment 2; N = narrative text; RD = reading disability; NR = not reported; T3 = Treatment 3; T4/C = Treatment group from a multiple treatment study coded as the contrast of interest; C = comparison; TRAVEL = write down the topic and circle it, read, verify the main idea by circling it and linking details, examine the next paragraph, and ask and verify again, link all the circles when finished with the story; R-dif = reading difficulty; A-R = at risk; TWA = thing before, while, and after reading; RAP = read, ask yourself what were the main ideas and details in the paragraph, put the main ideas and details in your own words; PLANS = pick writing goals, list ways to meet goals, make notes, and sequence notes; T2/C = treatment group from a multiple treatment study coded as the contrast of interest.

Table 3. Single-Case Design Study Information.

Study	Design	*N*	Grade	Struggling reader	Text type	Phase descriptions	Total sessions	Total hours	Fidelity reported	WWC design standards
Hagaman, Casey, & Reid (2016)	Multiple baseline	7	6 and 7	R-dif, LD	E	Baseline: read text, recall text information, answer comprehension questions Intervention: SRSD model to teach TRAP (think about what you are going to read prior to RAP strategy)	3–5	2.25–3.75	Y	MwR
Hagaman & Reid (2008)	Multiple baseline	3	6	R-dif	NR	Baseline: read text, retell, answer comprehension questions Intervention: SRSD to teach RAP strategy	4–5	2–2.5	Y	MwR
Hagaman, Casey, & Reid (2012)	Multiple baseline	6	9 and 10	R-dif	N, E	Baseline: read text, retell, answer comprehension questions Intervention: SRSD to teach RAP strategy	3–7	1–3.5	Y	MwR
Jitendra, Cole, Hoppes, & Wilson (1998)	Multiple baseline	3	6	LD	N	Baseline: read text, answer comprehension questions (multiple choice and main idea sentence generation) Intervention: Rules for identifying unrelated details, the main person or group, and the central action in a paragraph; cue card to monitor strategy use	10	6.7–8.3	Y	MwR
Lauterbach & Bender (1995)	Multiple baseline	2	9	LD	NR	Baseline: Brigance Diagnostic Inventory of Basic Skills reading comprehension probe. Intervention: Implement RAP with five-paragraph readings; cue card initially available for self-monitoring but faded by end of the phase	13, 15	NR	N	DNM
Mason, Snyder, Sukhram, & Kedem (2006)	Multiple baseline	9	4	R-dif, LD	E	Baseline: Read text, followed by oral and written retell assessment. Intervention: SRSD to teach TWA + PLANS strategy. After reading, students used color coding and a graphic organizer to identify main idea and details; orally rehearsed, and then wrote summaries.	15	7.5	Y	MwR

Note. WWC = What Works Clearinghouse; R-dif = reading difficulty; LD = learning disability; E = expository text; SRSD = self-regulated strategy development model; RAP strategy = read the text, ask yourself what was the main idea and two details, paraphrase what you read; MwR = meets design standards with reservation; NR = not reported; N = narrative text; DNM = does not meet design standards; TWA = think before, while, and after reading; PLANS = pick writing goals, list ways to meet goals, make notes, and sequence notes.

Table 4. Moderator Analyses by Year, Grade, Sessions, and Group Size.

	Number of effect sizes	Mean effect size	*SE*	95% CI	*p* value
Year (continuous)	103				
Intercept		0.97	0.14	[0.67, 1.26]	<.01
Slope		–0.01	0.01	[−0.04, 0.03]	.63
Grade	103				
Elementary (3–5)	35	0.68	0.12	[0.39, 0.97]	<.01
Middle and high school (6–12)	68	1.12	0.19	[0.72, 1.53]	<.01
Difference		0.44	0.22	[−0.04, 0.91]	.07
Treatment sessions	88				
<12	56	0.80	0.18	[0.40, 1.20]	<.01
13 or more	32	1.02	0.22	[0.48, 1.56]	<.01
Difference		0.22	0.28	[−0.40, 0.80]	.49
Group size	50				
1–4	35	1.41	0.30	[0.64, 2.19]	.01
5 or more	15	1.01	0.33	[0.11, 1.91]	.04
Difference		−0.40	0.44	[−1.43, 0.57]	.35

Note. CI = confidence interval.

Results From the Synthesis of SCD Studies

Five of the six SCD studies received a rating of *meets design standards with reservations*; at least one phase (i.e., baseline or intervention) in each study did not contain a minimum of five data points (Hagaman et al., 2012, 2016; Hagaman & Reid, 2008; Jitendra, Cole, Hoppes, & Wilson, 1998; Mason et al., 2006). The sixth study received a rating of *does not meet design standards* due to insufficient data in the baseline phase (i.e., only 1 data point), so we did not proceed to evidence standard evaluation for this study (Lauterbach & Bender, 1995). The results of the visual analysis (i.e., *strong, weak*, or *no evidence*), instructional practices addressed, and ES estimation (i.e., for the five studies that received a rating of *meets design standards with reservations*) are reported below.

Paraphrasing. Three studies examined the effect of the RAP paraphrasing strategy (i.e., read a paragraph, ask yourself what was the main idea and two details, put the main idea and details into your own words; Schumaker, Denton, & Deshler, 1994) on students' oral retell and accuracy with short-answer comprehension questions (Hagaman et al., 2012, 2016; Hagaman & Reid, 2008). In Hagaman and Reid (2008), visual analysis showed *strong evidence* in favor of the RAP strategy for oral retell and the short-answer comprehension measures. For oral retell, there was an immediacy of the effect from baseline to independent phase, no overlap, and an increase in level across three sixth-grade students. For short-answer comprehension questions, there was an immediate effect, no overlap, and a change in level for two students; the third student showed some overlap but an overall change in the level from baseline to independent phase (PND = 50; NAP = 92).

In Hagaman et al. (2012), RAP strategy instruction yielded *strong evidence* for oral retell, demonstrating an effect at three distinct time points across six ninth- and 10th-grade students. Visual analysis of the comprehension measure, however, yielded *no evidence* in favor of RAP strategy instruction. Two students (i.e., Brian and Katy) demonstrated an immediate effect, a change in level, and minimal overlap from the baseline to independent phase; however, this was not the case for the remaining four students who demonstrated unstable baseline performance, lack of an immediate effect, little to no change in level, or overlap between baseline and independent phases (PND = 40, 80, 0, 25; NAP = 73, 83, 68, 75). Consequently, the study did not demonstrate at least three functional relations at three distinct time points.

Finally, Hagaman et al. (2016) investigated the RAP strategy but added a before-reading step: Think about what you are going to read. This encouraged students to preview the text, attend to important text features (e.g., title, headings, etc.), and connect the text topic with existing prior knowledge. Visual analysis showed *strong evidence* for oral retelling and *moderate evidence* for comprehension across sixth- and seventh-grade participants. Four of the six students (i.e., Neil, Sarah, Leah, Robert, and Cara) demonstrated an effect on the comprehension assessment. Bailey, however, performed at a variable level during baseline with no immediate effect after the introduction of the RAP strategy and overlapping data between the two phases (PND = 20; NAP = 86). Tom achieved a stable baseline and an immediate change in level during independent phase,

but overlapping data points and a decreasing trend in the independent phase demonstrated a noneffect (PND = 20; NAP = 82).

Summarizing or main idea instruction with self-monitoring. Two studies examined main idea instruction using self-monitoring tools (Jitendra et al., 1998; Mason et al., 2006). In Jitendra et al. (1998), visual analysis showed an effect for two sixth-grade students (i.e., Chris and Tanya) on measures of narrative and expository comprehension but a noneffect for the third student. Brian's baseline data lacked stability, and there was no immediate effect from baseline to postintervention phase (PND = 40, 50, and NAP = 80, 67, on narrative and comprehension measures, respectively). The study did not demonstrate three effects at three distinct time points, resulting in *no evidence* for both measures.

Mason and colleagues (2006) used the RAP strategy with Brown and Day's (1983) five summarization rules to identify main ideas and develop summaries for each paragraph. Students used the PLANS strategy (i.e., pick goals for an essay, list ways to meet goals, make notes, and sequence the notes) to set and evaluate summary writing goals. The study did not meet design standards for oral retell as fewer than three data points were reported in the postinstruction phase for two of the three instructional groups. The written retell results, however, yielded *strong evidence* in favor of the intervention across three groups and nine students. All students demonstrated stable baseline data, an immediate effect, and an increase in level.

Discussion

The current review provided a meta-analysis (i.e., group design studies) and a synthesis (i.e., SCD studies) of the effects of main idea and summarizing interventions on the reading comprehension outcomes of struggling readers in Grades 3 through 12 from 1978 to 2016.

The Effectiveness of Main Idea and Summarizing Interventions

Group design studies. The mean estimate (ES = 0.97) indicates a positive effect of almost 1 standard deviation for main idea and summarizing interventions on struggling readers' reading comprehension. Although this suggests that main idea and summarizing interventions improve struggling readers' reading comprehension, the mean estimate reflects primarily unstandardized measures. Of the 103 ES estimates reported in 22 experimental studies, only seven effects resulted from standardized measures used in three experimental studies (i.e., AIMSweb Maze, Gates MacGinitie Reading Comprehension Test, Stanford Diagnostic Reading Test, Test of Reading Comprehension–3; Boyle, 1996; Faggella-Luby & Wardwell, 2011; Mason, 2013). Results from these measures yielded primarily negative, small effects and positive, small effects. This is consistent with previous reviews, demonstrating that researcher-developed measures are often associated with higher ESs than standardized measures (e.g., Edmonds et al., 2009). These results indicate that students' performance on more proximal measures of main idea and summarization is consistently better than on more generalized measures of reading comprehension. This finding may suggest that students are acquiring proficiency in the tasks they are taught; however, there is considerably less evidence that these effects generalize to broader measures of comprehension. As such, the results should be interpreted cautiously as standardized measures may be considered a more reliable and valid representation of an intervention's effectiveness.

Intervention intensity. We hypothesized that students would benefit more from interventions provided in smaller groups for an extended number of sessions compared to interventions provided in larger groups with fewer sessions. Our findings did not support this hypothesis as group size and number of sessions were not statistically significant predictors of ES. The results suggest that struggling readers benefit from main idea and summarizing interventions regardless of the group size (one to four vs. five or more) or number of sessions (12 or fewer vs. 13 or more). As with the grade-level analyses, it is important to mention that studies provided in smaller groups of one to four students result in a mean effect of almost 1½ standard deviations compared to a mean effect of 1 standard deviation for interventions provided in groups of five or more. Although both group sizes resulted in improved reading comprehension for struggling readers, it may be that small-group instruction allows for more targeted, individualized instruction and feedback. Previous research found that small-group intervention is more impactful in the elementary grades (e.g., Elbaum, Vaughn, Hughes, & Moody, 2000) but may not be a relevant intervention characteristic in the middle and high school grades (e.g., Vaughn et al., 2010). In this corpus, six studies provided intervention in small groups (i.e., one–four students), five studies provided intervention in groups of five or more, but almost half of the studies (i.e., 46%) did not report group size. This finding may differ had all studies reported group size. Further research is needed to determine the best ways to intensify summarizing and main idea interventions for struggling readers in Grades 3 through 12.

Grade level. The moderator analyses found no difference in the effectiveness of summarizing and main idea interventions for students in elementary versus middle and high school grades. This finding aligns with previous meta-analyses reporting no differences in intervention effectiveness by grade level (e.g., Wanzek et al., 2013). This indicates that students in the upper elementary, middle, and high

school grades benefit from summarizing and main idea practices. It is important to note that the mean estimate for middle and high school students resulted in a large effect of 1.12 compared to a moderate effect for elementary students of 0.68. It may be that middle and high school readers are better able to benefit from the types of instructional practices associated with summarization and main idea due to increased reading proficiency.

Publication year. We did not detect differences in ES as a function of publication year. Previous meta-analyses (e.g., Scammacca et al., 2015) report smaller mean ESs for studies published more recently. This decrease in ESs may be explained as a function of more rigorous research designs, an increased use of standardized measure, and improvements in the comparison condition (i.e., interventions tested in more recent studies are compared with school-provided interventions, whereas the comparison condition in older studies may not have received any intervention) since the establishment of the Institute for Education Sciences' WWC (Institute of Education Sciences, 2011). WWC aims to identify high-quality, evidence-based practices via rigorous design standards, thus providing guidance to the field regarding research-based practices. The majority of the studies included in this review were conducted prior to the establishment of the WWC. Five of the experimental studies and four SCD studies were conducted after 2002; none has been evaluated by WWC. Only six studies had investigated main idea or summarization instruction since 2010 ($n = 3$ group design studies; $n = 3$ SCD studies). It may also be that we were unable to detect differences in effects by publication year due to the overwhelming use of unstandardized measures or an insufficient number of more recent studies (i.e., only 26% of the group design studies were conducted after the inception of WWC).

Single-case research. The SCD studies, which addressed paraphrasing and self-monitoring practices, resulted in strong evidence in favor of main idea and summarizing interventions on measures of oral and written retell. These results support the findings from the group design studies, suggesting that paraphrasing with self-monitoring may enhance struggling readers' comprehension across the grade levels. However, results were less conclusive on short-answer comprehension measures (i.e., WWC evidence standard ratings ranged from *no evidence* to st*rong evidence*). As with the group design studies, the use of unstandardized measures (i.e., research-developed open-ended, short-answer comprehension questions) leaves many questions unanswered regarding the impact of summarizing and main idea interventions on generalized reading comprehension skill. As many of these measures were developed by the researchers for their respective studies, potential challenges related to measurement error may be related to student outcomes. More rigorous SCD studies (i.e., those that meet WWC standards without reservation) and improved validation of measures are needed.

Implications for Practice and Future Research

In the era of Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) and other state literacy standards, students are expected to identify main ideas and summarize texts; yet, struggling readers have great difficulty mastering these skills. Monitoring for meaning and determining the most important information present challenges for struggling readers. This review provides guidance related to promising practices practitioners might consider to support this population. For example, struggling readers may benefit from explicit instruction in narrative text structure (i.e., characters, setting, problem, events, solutions) and expository text structures (e.g., description, compare/contrast, sequence, cause/effect, problem/solution, etc.). In this review, text structure instruction improved students' recognition of text structure and text recall (Bakken et al., 1997; Dimino et al., 1990; Miller et al., 2011; Weisberg & Balajthy, 1989). Attending to the text's organization may help students situate ideas across paragraphs within the overall text structure; it may also facilitate retention of important information (Duke & Pearson, 2008). Another promising practice is paraphrasing, a micro-based approach in which students stop reading, reflect on the idea units presented across sentences in the text, and succinctly identify the most important information. Paraphrasing improved struggling readers' main idea identification and recall across grade levels (Ellis & Graves, 1990; Hagaman et al., 2012, 2016; Hagaman & Reid, 2008; Jenkins et al., 1987; Jitendra et al., 1998; Katims & Harris, 1997; Mason, 2004; Mason et al., 2006; Mastropieri et al., 2003). Teachers might consider pairing text structure and paraphrasing instruction with self-monitoring tools (e.g., cue card checklist, self-questioning) to support active engagement with text and facilitate monitoring for understanding.

Future research might examine practices that could further enhance the effectiveness of summarization and main idea interventions. For example, students may benefit from text structure and paraphrasing instruction. Combining text structure and paraphrasing may enable students to (a) identify the text's organizational pattern (e.g., compare/contrast) and (b) use an explicit paraphrasing process to identify the overarching main idea within that structure.

Another future research question might consider the role of systematic practice with high-quality feedback. Perhaps students' mastery of main idea and summarizing can be improved through repeated practice with process-specific feedback (Hattie & Timperley, 2007). Schunk and Rice (1992) investigated the effects of strategy-value feedback

(e.g., You did well because you followed the steps in the strategy); however, the effects of process- and product-specific feedback during main idea or summarizing interventions remain unknown. Researchers might consider embedding process-specific feedback to support students in tailoring their approach based on available text features (e.g., headings, keywords). Process-specific feedback may guide students in applying "fix-up practices" (i.e., a set of actions students choose from to address difficulty with a particular task) during the paraphrasing process. Potentially promising fix-up practices for identifying the main idea of a paragraph include checking the topic or concluding sentence (Katims & Harris, 1997; Miller et al., 2011), using self-monitoring prompt cards (Bakken et al., 1997; Boyle, 1996; Wong & Jones, 1982), color coding (Weisberg & Balajthy, 1990), rereading the text (Jenkins et al., 1987), and considering what might make a good title (Jenkins et al., 1987). Other fix-up practices may include checking the headings to provide a clue regarding that section's main idea or verifying one's main idea statement with self-questioning: (a) Is this the most important event that happened in this paragraph (i.e., narrative)? (b) Does the whole paragraph tell about this idea (i.e., expository)? Process-specific feedback may support struggling readers' skill in flexibly applying fix-up practices across text types.

Limitations

Although the findings are promising for improving students' reading comprehension, the results are limited in several ways. As mentioned previously, the results of this meta-analysis largely reflect findings from unstandardized measures, which are associated with larger ES estimates. A mean effect of standardized estimates would likely be lower than the mean effect presented here. The findings may be limited by the lack of unpublished studies included in this review. In spite of this, the Trim and Fill analysis indicated that no studies were missing as a result of publication bias. Finally, due to the time span (1978–2016) of studies included, the influence of the counterfactual must be taken into account when considering the findings. As a result of increased literacy standards in the last decade, it may be that teaching summarization and main idea is more prevalent now that it was 20 years ago, thus the counterfactual in the last 10 years may represent some overlap in instructional practices with the treatment. In some cases, the authors described the instruction in the comparison conditions (e.g., typical special education resource instruction or Tier 2 intervention; Faggella-Luby & Wardwell, 2011; Jitendra et al., 2000); however, many studies reported limited information on the comparison group. Limited or no instruction in the counterfactual likely inflates ES estimates, and limits the conclusions that can be made regarding summarizing and main idea intervention when compared with typical Tier 2 or Tier 3 interventions.

Conclusion

In sum, this review supports summarizing and main idea intervention as an effective practice for improving struggling readers' reading comprehension. Further research of high quality and rigor (i.e., fidelity of implementation, use of standardized measures, nature of the counterfactual's instruction) is needed to inform future practice and research by identifying high-impact instructional practices. In particular, this review provides support for further investigation of the effects of main idea and summarizing instruction using text structure and paraphrasing with repeated opportunities for practice. Future research might also investigate the effects of such instruction with and without targeted, process-specific feedback on the application of various fix-up practices to improve main idea generation.

Authors' Note

The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Declaration of Conflicting Interests

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

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported in part by the Institute of Education Sciences, U.S. Department of Education, through Grant R305F100013 to The University of Texas at Austin as part of the Reading for Understanding Research Initiative.

References

*Studies included in the systematic review.

Abdi, H. (2007). Bonferroni and Šidák corrections for multiple comparisons. In N. J. Salkind (Ed.), *Encyclopedia of measurement and statistics* (Vol. 3, pp. 103–107). Thousand Oaks, CA: SAGE.

*Bakken, J. P., Mastropieri, M. A., & Scruggs, T. E. (1997). Reading comprehension of expository science material and students with learning disabilities: A comparison of strategies. *The Journal of Special Education*, *31*, 300–324. doi:10.1177/002246699703100302

Biancarosa, C., & Snow, C. E. (2006). *Reading next—A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York* (2nd ed.). Washington, DC: Alliance for Excellent Education.

Borkowski, J. G., Weyhing, R. S., Carr, M. (1988). Effects of attributional retraining on strategy-based reading comprehension in learning-disabled students. *Journal of Educational Psychology*, *80*, 46–53. doi:10.1037/0022-0663.80.1.46

*Boyle, J. R. (1996). The effects of a cognitive mapping strategy on the literal and inferential comprehension of students with mild disabilities. *Learning Disability Quarterly*, *19*, 86–98. doi:10.2307/1511250

*Boyle, J. R., & Weishaar, M. (1997). The effects of expert-generated versus student-generated cognitive organizers on the reading comprehension of students with learning disabilities. *Learning Disabilities Research & Practice*, *12*, 228–235.

Brown, A. L., & Day, J. D. (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior*, *22*, 1–14. doi:10.1016/S0022-5371(83)80002-4

Chall, J. S. (1983). *Learning to read: The great debate*. New York, NY: McGraw-Hill.

*Dimino, J., Gersten, R., Carnine, D., & Blake, G. (1990). Story grammar: An approach for promoting at-risk secondary students' comprehension of literature. *The Elementary School Journal*, *91*, 19–32. doi:10.1086/461635

Duke, N. K., & Pearson, P. D. (2008). Effective practices for developing reading comprehension. *The Journal of Education*, *189*(1/2), 107–122.

Durkin, D. (1978). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly*, *14*, 481–533.

Duval, S., & Tweedie, R. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, *95*, 89–98. doi:10.1080/01621459.2000.10473905

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research*, *79*, 262–300. doi:10.3102/0034654308325998

Elbaum, B., Vaughn, S., Hughes, M., & Moody, S. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology*, *92*, 605–619. doi:10.1037//0022-0663.92.4.605

*Ellis, E. S., & Graves, A. W. (1990). Teaching rural students with learning disabilities: A paraphrasing strategy to increase comprehension of main ideas. *Rural Special Education Quarterly*, *10*, 2–10.

*Faggella-Luby, M., & Wardwell, M. (2011). RTI in a middle school: Findings and practical implications of a tier 2 reading comprehension study. *Learning Disability Quarterly*, *34*, 35–49. doi:10.1177/073194871103400103

Fisher, Z., Tipton, E., & Zhipeng, H. (2017). *Robumeta: Robust variance meta-regression* (R Package Version, 2.0). Retrieved from https://CRAN.R-project.org/package=robumeta

*Gajria, M., & Salvia, J. (1992). The effects of summarization instruction on text comprehension of students with learning disabilities. *Exceptional Children*, *58*, 508–516. doi:10.1177/001440299205800605

*Gallini, J. K., Spires, H. A., Terry, S., & Gleaton, J. (1993). The influence of macro and micro-level cognitive strategies training on text learning. *Journal of Research and Development in Education*, *26*, 164–178.

Goldman, S. R. (2012). Adolescent literacy: Learning and understanding content. *The Future of Children*, *22*, 89–106. doi:10.1353/foc.2012.0011

*Graves, A. W. (1986). Effects of direct instruction and metacomprehension training on finding main ideas. *Learning Disabilities Research*, *1*, 90–100.

Gwet, K. (2001). *Handbook of inter-rater reliability: How to estimate the level of agreement between two or multiple raters*. Gaithersburg, MD: STATAXIS Publishing Company.

*Hagaman, J. L., Casey, K. J., & Reid, R. (2012). The effects of the paraphrasing strategy on the reading comprehension of young students. *Remedial and Special Education*, *33*, 110–123. doi:10.1177/0741932510364548

*Hagaman, J. L., Casey, K. J., & Reid, R. (2016). Paraphrasing strategy instruction for struggling readers. *Preventing School Failure: Alternative Education for Children and Youth*, *60*, 43–52. doi:10.1080/1045988X.2014.966802

*Hagaman, J. L., & Reid, R. (2008). The effects of the paraphrasing strategy on the reading comprehension of middle school students at risk for failure in reading. *Remedial and Special Education*, *29*, 222–234. doi:10.1177/0741932507311638

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*, 81–112. doi:10.3102/003465430298487

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*, 39–65. doi:10.1002/jrsm.5

Institute of Education Sciences. (2011). *What works clearinghouse procedures and standards handbook* (Version 3.0). Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf

*Jenkins, J. R., Heliotis, J. D., Stein, M. L., & Haynes, M. C. (1987). Improving reading comprehension by using paragraph restatements. *Exceptional Children*, *54*, 54–59.

Jitendra, A. K., Chard, D., Hoppes, M. K., Renouf, K., & Gardill, M. C. (2001). An evaluation of main idea strategy instruction in four commercial reading programs: Implications for students with learning problems. *Reading & Writing Quarterly*, *17*, 53–73. doi:10.1080/105735601455738

*Jitendra, A. K., Cole, C. L., Hoppes, M. K., & Wilson, B. (1998). Effects of a direct instruction main idea summarization program and self-monitoring on reading comprehension of middle school students with learning disabilities. *Reading & Writing Quarterly*, *14*, 379–396. doi:10.1080/1057356980140403

*Jitendra, A. K., Hoppes, M. K., & Xin, Y. P. (2000). Enhancing main idea comprehension for students with learning problems: The role of a summarization strategy and self-monitoring instruction. *The Journal of Special Education*, *34*, 127–139. doi:10.1177/002246690003400302

Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., & Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A practice guide* (NCEE #2008-4027). Washington, DC: 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

*Katims, D. S., & Harris, S. (1997). Improving the reading comprehension of middle school students in inclusive classrooms. *Journal of Adolescent & Adult Literacy*, *41*, 116–123.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*, 363–394. doi:10.1037/0033-295X.85.5.363

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010).

Single-case designs technical documentation. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_scd.pdf

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education*, *34*, 26–38. doi:10.1177/0741932512452794

*Lauterbach, S. L., & Bender, W. N. (1995). Cognitive strategy instruction for reading comprehension: A success for high school freshmen. *The High School Journal*, *79*, 58–64.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis* (Vol. 49). Thousand Oaks, CA: SAGE.

*Malone, L. D., & Mastropieri, M. A. (1992). Reading comprehension instruction: Summarization and self-monitoring training for students with learning disabilities. *Exceptional Children*, *58*, 270–279. doi:10.1177/001440299105800309

*Mason, L. H. (2004). Explicit self-regulated strategy development versus reciprocal questioning: Effects on expository reading comprehension among struggling readers. *Journal of Educational Psychology*, *96*, 283–296. doi:10.1037/0022-0663.96.2.283

*Mason, L. H. (2013). Teaching students who struggle with learning to think before, while, and after reading: Effects of self-regulated strategy development instruction. *Reading & Writing Quarterly*, *29*, 124–144. doi:10.1080/10573569.2013.758561

*Mason, L. H., Snyder, K. H., Sukhram, D. P., & Kedem, Y. (2006). TWA + PLANS strategies for expository reading and writing: Effects for nine fourth-grade students. *Exceptional Children*, *73*, 69–89. doi:10.1177/001440290607300104

*Mastropieri, M. A., Scruggs, T. E., Spencer, V., & Fontana, J. (2003). Promoting success in high school world history: Peer tutoring versus guided notes. *Learning Disabilities Research & Practice*, *18*, 52–65. doi:10.1111/1540-5826.00057

Meyer, B. J. F., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly*, *16*, 72–103. doi:10.2307/747349

*Miller, C. A., Darch, C. B., Flores, M. M., Shippen, M. E., & Hinton, V. (2011). Main idea identification with students with mild intellectual disabilities and specific learning disabilities: A comparison of explicit and basal instructional approaches. *Journal of Direct Instruction*, *11*, 15–29.

Moeyaert, M., Maggin, D., & Verkuilen, J. (2016). Reliability, validity, and usability of data extraction programs for single-case research designs. *Behavior Modification*, *40*, 874–900. doi:10.1177/0145445516645763

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(6), e1000097. doi:10.1371/journal.pmed1000097

National Center for Education Statistics. (2015). *The nation's report card: Trends in academic progress 2015* (NCES 2015–136). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards English language arts*. Washington, DC: Author.

Newman, L., Wagner, M., Huang, T., Shaver, D., Knokey, A.-M., Yu, J., . . . Cameto, R. (2011). *Secondary school programs and performance of students with disabilities: A special topic report of findings from the National Longitudinal Transition Study-2 (NLTS2)* (NCSER 2012–3000). Washington, DC: U.S. Department of Education, National Center for Special Education Research.

Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy*, *40*, 357–367. doi:10.1016/j.beth.2008.10.006

Parker, R. I., Vannest, K., & Davis, J. (2014). Non-overlap analysis for single-case research. In T. R. Kratochwill & J. Levin (Eds.), *Single-case intervention research* (pp. 127–151). Washington, DC: American Psychological Association.

Pustejovsky, J. E. (2017). *Single-case effect size calculator* (Version 0.2) Web application. Retrieved from https://jepusto.shinyapps.io/SCD-effect-sizes/

Rapp, D. N., van den Broek, P., McMaster, K. L., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading*, *11*, 289–312. doi:10.1080/10888430701530417

Rohatgi, A. (2015). *WebPlotDigitizer user manual version 3.9*. Retrieved from https://automeris.io/WebPlotDigitizer/userManual.pdf

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

Schumaker, J. B., Denton, P. H., & Deshler, D. D. (1994). *The paraphrasing strategy: Instructor's manual*. Lawrence, KS: University of Kansas Institute for Research in Learning Disabilities.

*Schunk, D. H., & Rice, J. M. (1992). Influence of reading-comprehension strategy information on children's achievement outcomes. *Learning Disability Quarterly*, *15*, 51–64. doi:10.2307/1510565

Scruggs, T. E., & Mastropieri, M. A. (1998). Summarizing single-subject research: Issues and applications. *Behavior Modification*, *22*, 221–242. doi:10.1177/01454455980223001

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education*, *8*, 24–33. doi:10.1177/074193258700800206

Solis, M., Ciullo, S., Vaughn, S., Pyle, N., Hassaram, B., & Leroux, A. (2012). Reading comprehension interventions for middle school students with learning disabilities: A synthesis of 30 years of research. *Journal of Learning Disabilities*, *45*, 327–340. doi:10.1177/0022219411402691

Taylor, K. K. (1986). Summary writing by young children. *Reading Research Quarterly*, *21*, 193–208. doi:10.2307/747845

Tipton, E. (2013). Robust variance estimation in meta-regression with binary dependent effects. *Research Synthesis Methods*, *4*, 169–187. doi:10.1002/jrsm.1070

Valentine, J. C., & Cooper, H. (2008). A systematic and transparent approach for assessing the methodological quality of intervention effectiveness research: The Study Design and Implementation

Assessment Device (Study DIAD). *Psychological Methods*, *13*, 130–149. doi:10.1037/1082-989X.13.2.130

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

Vaughn, S., Klingner, J. K., Swanson, E. A., Boardman, A. G., Roberts, G., Mohammed, S. S., . . . Stillman-Spisak, S. J. (2011). Efficacy of collaborative strategic reading with middle school students. *American Educational Research Journal*, *48*, 938–964. doi:10.3102/0002831211410305

Vaughn, S., Wanzek, J., Wexler, J., Barth, A., Cirino, P. T., Fletcher, J., . . . Francis, D. (2010). The relative effects of group size on reading progress of older students with reading difficulties. *Reading and Writing*, *23*, 931–956. doi:10.1007/s11145-009-9183-9

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*, 163–195. doi:10.3102/0034654313477212

Watson, S. M. R., Gable, R. A., Gear, S. B., & Hughes, K. C. (2012). Evidence-based strategies for improving the reading comprehension of secondary students: Implications for students with learning disabilities. *Learning Disabilities Research & Practice*, *27*, 79–89. doi:10.1111/j.1540-5826.2012.00353.x

*Weisberg, R., & Balajthy, E. (1989). Transfer effects of instructing poor readers to recognize expository text structure. *National Reading Conference Yearbook*, *38*, 279–286.

*Weisberg, R. K., & Balajthy, E. (1990). Development of disabled readers' metacomprehension ability through summarization training using expository text: Results of three studies. *Journal of Reading, Writing, and Learning Disabilities International*, *6*, 117–136. doi:10.1080/0748763900060204

Williams, J. P. (2006). Stories, studies, and suggestions about reading. *Scientific Studies of Reading*, *10*, 121–142. doi:10.1207/s1532799xssr1002_1

*Wong, B. Y., & Jones, W. (1982). Increasing metacomprehension in learning disabled and normally achieving students through self-questioning training. *Learning Disability Quarterly*, *5*, 228–240. doi:10.2307/1510290