SELF-QUESTIONING STRATEGY FOR STRUGGLING READERS: A SYNTHESIS

Johny Daniel, M.Ed.[1]

Kelly J. Williams, Ph.D.

Author Note

Johny Daniel, Department of Special Education, University of Texas at Austin

Kelly J. Williams, Department of Curriculum and Instruction, Indiana University, Bloomington

Correspondence

[1]University of Texas at Austin

College of Education, SZB 228

1912 Speedway D4900

Austin, TX – 78712

Johny.daniel@utexas.edu

Abstract

This review examined the effects of self-questioning (SQ) strategy instruction on reading comprehension outcomes for students with learning disabilities and struggling readers in grades K-12. Our literature search, encompassing the past 53 years (1965-2018) of research, found ten studies that fit our inclusion criteria. Reviewed studies included eight group-design and two single-case design studies. Overall, the effects of SQ strategy instruction on students' reading comprehension outcomes were mixed. No clear trends of the effects of SQ strategy intervention were associated with participants' grade-level and type of instruction (explicit or nonexplicit instruction). Effects of the total number of hours of SQ strategy instruction on students' reading outcomes varied slightly with medium to large effects for students receiving two or more total hours of strategy instruction.

Self-Questioning Strategy for Struggling Readers: A Synthesis

Reading comprehension is defined as the process of extracting and constructing meaning while actively engaging with text (Snow, 2002). Students who are successful at reading and comprehending grade-level texts actively monitor their reading by using a set of reading comprehension strategies to extract and construct meaning (Baker & Brown, 1984; Pressley & Afflerbach, 1995; Pressley, Borkowski, & Schneider, 1989; Snow, 2002; Wilkinson & Son, 2011). On the other hand, some students who struggle with reading comprehension neglect to establish a purpose for reading, inadequately monitor comprehension, and do not make connections within text (Paris, Wasik, & Turner, 1991; Wilkinson & Son, 2011).

Reading Comprehension Strategies

Successful readers use strategies to monitor thoughts and ideas about learning that facilitate self-assessment of text comprehension. In the context of reading comprehension, the term 'strategy' refers to a cognitive or behavioral action that is performed to allow readers to evaluate some aspect of knowledge from the text being read (e.g., main idea, inference making, factual information) (Graesser, 2012). Research related to reading comprehension intervention has shown that acquiring and implementing reading comprehension strategies positively impacts students' reading comprehension outcomes (e.g., Berkeley, Scruggs, & Mastropieri, 2010; Gersten, Fuchs, Williams & Baker, 2001). Pressley and colleagues (1995) determined that the use of comprehension strategies allows struggling readers to emulate skilled readers and thus leads to improved comprehension. Therefore, the structure that reading comprehension strategies provide students for unstructured tasks such as reading may be the key contributing factor (Rosenshine, Meister, & Chapman, 1996). In addition, reading comprehension strategies

promote repair of comprehension failure during the reading process to improve text comprehension (Wagoner, 1983).

One example of a packaged set of strategies is Collaborative Strategic Reading ([CSR], Klingner & Vaughn, 1998). During CSR, students learn to use various strategies to monitor their reading comprehension. When reading text that is challenging to comprehend, students preview the text, find the main idea, generate questions to monitor understanding, and review key ideas. In this example, CSR is designed to provide learners ready access to a structural framework that scaffolds the reading task and provides a purpose for reading. When students approach a reading task with purpose—such as generating questions—they have opportunities to make cognitive judgements about which aspects of the text are most important (Baker & Brown, 1984). Notably, Rosenshine and colleagues (1996) suggest that using comprehension strategies, such as generating questions while reading, improves comprehension because they activate cognitive processes while searching text, making inferences, combining information, and monitoring understanding.

Of the many reading comprehension strategies, generating questions while reading is one of the most commonly recommended. For instance, according to the influential National Reading Panel (NRP) report, teaching students to generate questions while reading was described to have the *strongest scientific evidence* to effectively improve students' text comprehension (National Reading Panel, 2000). Similarly, in his analysis of the NRP report, Willingham (2006-2007) noted that of all the NRP recommended effective reading comprehension strategies, conclusive evidence was found for two: self-questioning and multiple strategy instruction. Evidence-based multi-strategy packages such as Peer-Assisted Learning Strategy and CSR have question generation embedded as part of multi-component programs (Shanahan et al., 2010).

Additionally, other well-established reading strategies such as Know-Want to Know-Learn ([KWL]; Ogle, 1986), Survey-Question-Read-Recite-Review ([SQ3R]; Robinson, 1941), Read-Ask-Paraphrase ([RAP]; Schumaker, Denton, & Deshler, 1984), and Think-Read-Ask-Paraphrase ([TRAP]; Hagaman, Casey, & Reid, 2016) also teach question generation as part of multicomponent strategy instruction. Past syntheses (Rosenshine et al., 1996; Wong, 1985) on self-questioning strategy instruction have shown benefits, whether used as part of multicomponent interventions or as an isolated strategy intervention, on reading comprehension outcomes for typically developing students. However, the nature of association between self-questioning instruction as a singular reading comprehension strategy and its effects on improving reading comprehension outcomes for struggling readers including students with learning disabilities (LD) is less clear.

While self-questioning strategy instruction has been recommended to improve readers' level of comprehension (National Reading Panel, 2000; Willingham, 2006-2007), no previous synthesis has studied the efficacy of self-questioning strategy instruction on struggling readers' comprehension outcomes. Therefore, the purpose of this study is to evaluate the effectiveness of self-questioning strategy instruction on reading comprehension outcomes for struggling readers in grades K-12.

Self-Questioning Strategy

Thorndike (1917) reasoned that "the vice of a poor reader is to say the words to himself without actively making judgements concerning what they reveal" (p.332). In other words, struggling readers generally read on auto-pilot mode without questioning their understanding of the text being read (Duffy & Roehler, 1987) nor the authority of the written text (Paris et al., 1991). In contrast, asking content related questions while reading is characteristic of strategic

readers (Collins, Brown, & Larkin, 1980). The fundamental idea of a self-questioning strategy is for students to generate questions about the text they are reading to help monitor their understanding, typically through regularly stopping and asking themselves questions (Andre & Anderson, 1978; Frase & Schwartz, 1975). Thus, students think about the content and actively engage with the task. Furthermore, it affords them the ability to stop when they are unable to answer self-generated questions and review the text or activate their background knowledge before proceeding to complete the reading task (Mastropieri & Scruggs, 1997; Wong, 1985).

Over the last few decades, researchers have investigated different procedural paths to improve the effectiveness of self-questioning strategy use for all students (Rosenshine et al., 1996; Wong, 1985). Two such paths include the top-down and bottom-up approach. The top-down approach refers to what readers bring to the reading task. In this approach readers are instructed in the use of self-questioning strategy before they read a text. They then use their knowledge of the strategy to generate and answer their own questions during or after the reading activity (e.g., Holmes, 1985; Rouse, Alber-Morgan, Cullen, & Swayer, 2014). Alternately, the bottom-up approach refers to teachers determining the aspects of the text students need to focus on. In this approach, teachers provide students with questions before the reading activity, and students answer teacher-generated questions during or after the reading task (e.g., Crabtree, Alber-Morgan, & Konrad, 2010; Taylor, Alber, & Walker, 2002).

Bottom-up approaches have shown to be effective in improving students' decoding skills (Shaywitz et al., 2004; Torgesen et al., 2001; Torgesen et al., 1999). However, bottom-up approaches do not adequately support higher-level skills needed to comprehend text (Foorman, Francis, Fletcher, Schatschneider, & Mehta 1998; Torgesen et al., 1999). To strengthen comprehension, top-down approaches are also required as they provide students with skills

needed to understand the logic and structure in the text (McCardle, Scarborough, & Catts, 2001; Nation, 2005). Although teacher-provided questions allow students to focus on answering given questions, they do not contribute to student generalization of this strategy in different contexts due to student dependence on teacher generated questions. In contrast, teaching students to independently use the strategy through a top-down approach, provides them with tools to problem solve comprehension failures independently. Hence, only interventions that used the top-down approach to learning self-questioning strategy were included in this synthesis.

Past self-questioning strategy instruction reviews. Since 1985, two syntheses have examined the effects of self-questioning strategy instruction on readers' comprehension (Rosenshine et al., 1996; Wong, 1985). Participants of both syntheses included individuals with and without disabilities. Furthermore, both syntheses included studies that implemented self-questioning strategy in conjunction with one or more reading comprehension strategies (e.g., summarization).

Of the 53 studies in both syntheses (Rosenshine et al., 1996, $n = 26$; Wong, 1985, $n = 27$), only eight studies included students with LD, or students identified as below-average/struggling readers, and manipulated self-questioning strategy as the sole independent variable. In four of the eight studies analyzed in both syntheses, students in the treatment group did not significantly differ from comparison group peers on researcher developed or standardized posttest reading measures (Bernstein, 1973; Simpson, 1989; Smith, 1972; Smith, 1977). In contrast, Cohen (1983), and Wong and Jones (1982) reported significant positive effects on comprehension outcomes for treatment group students. Davey and McBride (1986) reported significant positive effects on the posttest measure of inferential questions and nonsignificant differences on the posttest measure of literal comprehension questions.

More recently, Joseph, Alber-Morgan, Cullen, and Rouse (2016) conducted a literature review on the effects of self-questioning strategy on students' comprehension outcomes. The review evaluated results from 35 studies published between 1990–2012. The authors found sufficient evidence of the effectiveness of self-questioning strategy instruction on students' comprehension outcomes. However, similar to the previous two self-questioning syntheses (Rosenshine et al., 1996; Wong, 1985), Joseph and colleagues evaluated studies that taught students to use self-questioning strategy in isolation and also in combination with other reading strategies (e.g., KWL [Know, Want to know, What I learned]). Furthermore, studies included in the literature review included students with and without LD.

In summary, the two research syntheses and one literature review conducted on self-questioning strategy instruction included only a small percentage of studies involving students with LD or struggling readers and concurrently limited treatment solely to self-questioning strategy instruction. Of the eight studies included in the two past syntheses, three studies showed significant positive differences in reading comprehension outcomes for treatment group students. While research literature (e.g., Berkeley, Scruggs, & Mastropieri, 2010; National Reading Panel, 2000; Pearson & Fielding, 1991) widely cites the importance of self-questioning strategy instruction and advocates the use of this strategy for students, it is important to note that the benefits of self-questioning strategy for struggling readers is based on scant literature.

In contrast to past syntheses and literature review on the topic, this synthesis focuses solely on reading outcomes for a subpopulation of students who are identified with LD or as struggling readers in grades K-12. Another essential contrast to past reviews is that the current review aims to analyze the isolated effect of self-questioning strategy instruction (i.e., not combined with other reading strategies) on reading outcomes for struggling readers.

Research Question

1. Does self-questioning strategy instruction lead to improvements in reading

    comprehension for students with LD or struggling readers in grades K-12?

    a. What are the instructional components and conditions under which self-

       questioning strategy intervention studies are most effective in improving

       struggling readers' comprehension outcomes?

    b. Applying the What Works Clearinghouse (WWC) initial screening criteria

       (WWC, 2017), how many studies that implement self-questioning strategy

       instruction to improve struggling readers' comprehension outcomes are

       methodically sound?

Method

Operational Definition

In context of this article, struggling reader refers to a student who scored significantly

below their age- or grade-level peers on a measure of reading ability. The term struggling readers

also includes students identified with a LD, a reading disability/difficulty, or at-risk for reading

difficulty.

Data Collection

For this synthesis, an extensive search of the literature was conducted using a four-step

process. First, a computer search of peer reviewed articles was conducted on ERIC, PsycINFO,

and Education Source to locate studies published between 1965 and September 2018. Search

terms used (ERIC, Education Source, and PsycINFO) were as follows: Line 1: read*; Line 2:

"reading diff*" OR "learning diff*" OR "reading dis*" OR "learning dis*" OR "reading

problems" OR "learning problems" OR "struggling readers" OR dyslexia OR "reading delay" OR

"mild handi*" OR "special needs" OR "special education" OR "with disabilities" OR "at risk"

OR "high risk;" Line 3: question* OR SRSD OR "self regulat*" OR "self manage*" OR "self

monitor*" OR "self generated" OR summariz* OR "cognitive strategies" OR "main idea" OR

RAP OR paraphra* OR QAR OR QRAC OR KWL OR PQ4R OR SQ3R. Line 1 terms specified

the content area of instruction. Line 2 terms identified the population with whom the studies

were to be conducted (e.g., struggling readers, students with disabilities). Line 3 terms were used

to narrow down the specific strategies used in the study and included key words and acronyms

for common self-questioning strategies. For example, SQ3R (Robinson, 1941) is a specific

question generation strategy.

The second step in identifying articles relevant to the research question involved a hand

search of 14 prominent educational journals spanning from 2013 through September, 2018. A

five-year window was searched to ensure that the electronic search had captured all relevant

articles. As no additional articles were found in the hand search, we did not go back and search

in years prior to 2013. The following journals were hand searched: *Annals of Dyslexia, Cognition*

and Instruction, Exceptional Children, Journal of Educational Psychology, Journal of Learning

Disabilities, Journal of Special Education, Learning Disability Quarterly, Learning Disabilities

Research and Practice, Reading and Writing Quarterly, Reading Psychology, Reading Research

Quarterly, Remedial and Special Education, Review of Educational Research, and Scientific

Studies of Reading. Next, relevant articles were sourced via an ancestry search of articles that fit

the inclusion criteria. Finally, to avoid publication bias and address the 'file drawer' problem

(Rosenthal, 1979) that refers to studies that are unpublished due to null results, we contacted

primary investigators (researchers whose contact information could be obtained from the

Worldwide Web) of studies included in this synthesis to solicit information on research authors may have conducted on this topic but not published.

Studies that met the following criteria were included in the synthesis: (a) interventions involving participants identified with LD, dyslexia, or struggling readers in grades K-12 (at least 50% or more of sample if aggregated data), (b) randomized control trials, quasi-experimental designs, or single-case designs in which self-questioning strategy interventions were implemented to improve reading comprehension performance, (c) self-questioning strategy interventions that involved student generated questions, (d) studies that included at least one measure of reading comprehension, and (e) studies published and taught in English. Studies were excluded if they were not conducted in school settings, were multicomponent interventions (e.g., self-questioning and summarizing, or self-questioning and making predictions), or were literature reviews, meta-analyses, observational studies, news reports, conference proceedings, under review, or in-preparation articles.

For studies that included students with different disabilities, we only used data for students with LD, or students identified as struggling readers with no comorbid disability. For example, Berkeley et al. (2011) included a heterogenous population of students in their study that included typically developing students, students with hearing impairment, LD, and other health impairments. For the purpose of this systematic review, we contacted authors to gain access to disaggregated data for students identified with LD and only used data for those students in our analysis.

Coding Procedures

Online database search revealed 8,064 potential articles. The primary investigator screened articles using titles and abstracts to exclude studies not germane to the topic. After the

initial screening, 103 articles were identified for potential inclusion. The primary investigator reviewed full text for all 103 studies. Of the 103 studies, ten studies fit all inclusion criteria. The primary investigator coded all ten studies and then the secondary coder double-coded the original code-sheets to determine discrepancies. Any discrepancies in coding were discussed and authors mutually agreed on a solution. Initial interobserver agreement (IOA) was 96.15%, and post discussion IOA was 100%. A further breakdown of the IOA by response type is provided in Appendix A.

Studies were coded using the Guide for Education-Related Intervention Study Syntheses (Vaughn, Elbaum, Wanzek, Scammacca, & Walker, 2014). This codesheet has been used in previous educational syntheses (e.g., Hall, 2016; Hall et al., 2017; Scammacca et al., 2016; Stevens, Walker, & Vaughn, 2017; Williams, Walker, Wanzek, & Vaughn, 2017) and includes all critical components identified in the systematic review process of the WWC Study Review Guide (WWC, 2017). Critical components included in the codesheet are: design information, sample description; sample sizes; baseline measures; measures' description including validity, reliability and internal consistency information of each measure; data used for analysis; attrition information; description of treatment and control groups; and description of treatment and control group procedures.

Participant information. First, studies were coded for participant information. Items in this section were forced choice response type. Options in this section included information about participants' socio-economic status, risk type (e.g., at-risk, LD, dyslexia), student age, grade, and gender.

Design information. Detailed information about the study design were recorded on the code-sheet. Forced choice options in this section included the type of study design (e.g.,

treatment/comparison randomized experiment, treatment/comparison quasi-experiment, single-case design, single group design), the assignment of participants to groups (i.e., random, matched, not reported, or other) and whether implementation fidelity and pretest scores were reported. Additionally, one open-ended question included description of criteria for selection of participants.

Treatment/comparison groups. For each treatment and comparison group in each study, a table was completed. This table required entering data related to participants' age and grade-level, dosage of the intervention (i.e., duration, frequency, total number of sessions), and group size. One forced choice option recorded information on person implementing the intervention (e.g., teacher, researcher). Furthermore, one open-ended question included description of treatment and control group procedures.

Attrition. Forced choice options were used to record the sample sizes of the study groups at the beginning and end of the study. Using the WWC criteria for overall and differential attrition (WWC, 2017), the sample size information was used to determine if overall and/or differential attrition could have possibly impacted study results.

Study quality. During coding, studies were rated as *meets standards without reservations, meets standards with reservations,* or *does not meet standards*. Studies rated as *meets standards without reservations* met the WWC initial screening criteria (WWC, 2017) recommending that group membership be determined through a random process and attrition whether differential or overall be in the range of *low attrition*. Group design studies were rated as *meets standards with reservations* when baseline equivalence was reported and participants were randomly assigned but there was high overall or differential attrition. Group design studies were

rated as *does not meet standards* when there was no attempt to randomize and no baseline

equivalence reported to determine causality of results.

Similarly, for both single-case design studies quality ratings were based on WWC's

single-case design standards (WWC, 2017). We used the WWC criteria for evaluating single-

case design studies (WWC, 2017). Studies received a rating of *meets design standards* when the

following criteria were observed: (a) systematic manipulation of an independent variable (i.e.,

self-questioning strategy intervention); (b) more than one assessor systematically measured the

outcome variable(s) over a period of time with more than 0.80 IOA on at least 20% of the data

points; (c) at least three attempts to demonstrate intervention effects were systematically

measured at three different points in time to establish experimental control (i.e., minimum of

three baseline and three intervention phases in multiple baseline design); (d) each phase included

at least five data points. A study was rated *meets design standards with reservations* when it met

criteria a, b, c, and had at least three to four data points in each phase. A study was rated *does not*

meet design standards with or without reservations when a study did not meet criteria a, b, or c,

or reported less than three data points in each phase.

Study results. Forced choice options were used to examine for precision of outcome. It

was recorded whether or not the authors noted that statistical assumptions were or were not met

(i.e., assumption of normality, independence, equal variance). Next, raw data was extracted from

the original study (i.e., mean, standard deviation, sample size) and inputted into a table. This

information also included measure name, measure type (e.g. comprehension, fluency, etc.), min

and max score, and reliability and validity coefficients.

Data Analysis

Group design studies. Standardized mean difference effect sizes were calculated using

Hedges's *g* to adjust for the possibility of small sample bias. We used treatment and comparison

groups' posttest means, standard deviations, and sample sizes to calculate Hedges's *g*. We

considered all eligible effect sizes in each study that provided mean and standard deviation (i.e., 6

studies included) to calculate a weighted mean effect size; two group design studies were not

included in measuring the weighted mean effect size because one study (Chan, 1991) lacked a

control group and another study (Wong & Jones, 1982) did not report descriptive statistics for

reading outcomes. Group design studies could contribute multiple effect sizes as long as the

sample for each effect size was independent. For studies that reported multiple effect sizes from

the same sample (e.g., two effect sizes based on two reading comprehension measures were

calculated for treatment versus control in one study), we accounted for the statistical dependencies

using the random effects robust standard error estimation technique developed by Hedges, Tipton,

and Johnson (2010). This analysis allows for the clustered data (i.e., effect sizes nested within

samples) by correcting the study standard errors to take into account the correlations between

effect sizes from the same sample. The robust standard error technique requires that an estimate of

the mean correlation (ρ) between all the pairs of effect sizes within a cluster be estimated for

calculating the between-study sampling variance estimate, τ^2. In all analyses, we estimated τ^2 with

$\rho = .80$; sensitivity analyses showed that the findings were robust across different reasonable

estimates of ρ. Because we included studies conducted in kindergarten through 12th grade, we

hypothesized that the research body is reporting a distribution of effect sizes with significant

between-studies variance, as opposed to a group of studies attempting to estimate one true effect

size. Thus, we used a random-effect model for the current study (Lipsey &Wilson, 2001).

Cohen (1988) recommends interpreting effect sizes of .20 as "small," 0.50 as "medium,"

and 0.80 as "large," while the WWC recommends interpreting effect sizes of 0.25 and larger as

"substantially important" in educational research settings (WWC, 2017, p.D-5). We considered both of these interpretations when interpreting the magnitude and importance of the effects. Finally, we used descriptive statistical data to calculate 95% confidence intervals to determine if each individual effect size was significant; that is, if a statistic is significantly different from 0 at the 0.05 level, then the 95% confidence interval will not contain 0.

Single-case design studies. Two single-case design studies were included in this synthesis. For one study (Rouse et al., 2014), Tau-U was calculated for each participant to show the magnitude of difference between baseline phase data and the final phase data when self-questioning instruction was faded. Tau-U is a robust measure for effect size calculation for small data sets (Parker, Vannest, Davis, & Sauber, 2011; Vannest & Ninci, 2015). The Tau-U results were interpreted according to Vannest and Ninci (2015) categorization: 0.20 indicating a small change, 0.20 to 0.60 a moderate change, 0.60 to 0.80 a large change, and 0.80 or greater a very large change between two phases. We used an online web-application, Single-Case Effect Size Calculator, to measure nonoverlap and parametric effect size from single-case design study data and to calculate Tau-U (Pustejovsky, 2017). However, it was not possible to calculate Tau-U for Clark, Deshler, Schumaker, Alley and Warner, (1984) as the manuscript only showed a graph of different phases for one of the six study participants. The study reported a combined accuracy percentage for all six students but raw data for each phase and individual student graphs were not available to calculate Tau-U or conduct visual analysis.

In addition to calculating Tau-U, visual analysis procedures were used to measure the relationship between self-questioning strategy intervention and struggling readers' reading outcomes. As per WWC recommendations (2017), visual analysis was conducted to evaluate the effects within single-case design studies that were rated as *meets design standards with or*

without reservations. Visual analysis was conducted using the following procedures to evaluate

the within- and between-phase data for multiple baseline design studies:

1. Determined if baseline pattern is predictable and stable.

2. Examined within each phase the level, trend, and variability of data. Level refers to the

    average score in each phase, trend refers to the best-fitting slope line for data in each

    phase, and variability refers to the variability of data around the trend line in each phase.

3. Compared between similar/adjacent phases to assess whether manipulation of

    independent variable was associated with a predicted change in the dependent variable by

    comparing overlap, immediacy of effect, and consistency of data pattern.

4. Combined information from all phases to determine whether study reported a minimum

    of three demonstrations of an effect at three different points in time.

    Study findings were characterized as *strong evidence, moderate evidence*, or *no evidence*.

A study that reported three demonstrations of an effect at three different points in time received a

strong evidence rating; if a study reported one demonstration of a non-effect in addition to three

demonstrations of an effect it received a *moderate evidence* rating. Studies that did not report at

least three demonstrations of an effect received a *no evidence* rating.

Results

Ten[1] studies met inclusion criteria and were included in this synthesis. Nine were located

through the online database search and one through the ancillary search. Table 1 summarizes

study characteristics and table 2 summarizes findings of eight group and two single-case design

studies. A total of 474 students participated in interventions across the ten included studies. For

---

[1] One primary investigator, who was contacted to address the "file drawer problem," sent us a dissertation on the topic. We did not include this study as it did not meet our search criteria.

the current study's analysis, we included results for 129 students identified as students with LD/reading disability (RD) and 137 students identified as struggling readers. The remainder of the sample was excluded for the following reasons: participant was a typically-developing student, participant had a disability other than LD/RD, and/or participant was randomized to a group that taught self-questioning strategy in combination with another strategy or taught another strategy.

Research Question 1: Effectiveness of Self-questioning Strategy Instruction

The analysis showed that for group design studies the estimated average weighted effect size between the treatment and control groups on reading outcomes was $g = .61$, CI 95% [.30, .92], ($\tau^2 = 0.14$). Because of the small number of studies, we could not run meta-regression to explore moderators. Statistical significance was also measured for each effect size. Of the sixteen Hedge's g effect sizes measured, six were positive and significant while ten were not significant.

In both single case design studies, participants scored approximately 30 to 50 percentage points higher in the treatment phase compared to baseline phase on researcher developed measures of reading comprehension. For one single-case design study, raw data were available to calculate an effect size using Tau-U; large effects were computed for both participants in the study (Min = 1.00, Max = 1.02) (Rouse et al., 2014). However, it was not possible to conduct visual analysis for both single-case design studies as they did not meet WWC design standards due to the lack of three demonstration of intervention effects at three different points in time and/or lack of three or more data points in each phase.

Research Question 1a: Instructional Components and Conditions Under Which Self-Questioning Strategy Instruction is Most Effective for Struggling Readers in grades K-12

  Participants.

Identification. Each study used different methods to identify struggling readers. Two studies (Davey & McBride, 1986; Holmes, 1985) used the California Achievement Test (CTB/McGraw-Hill, 1978) data to identify participants' grade level placement for reading. Additionally, Holmes (1985) solicited teacher recommendations on students' reading proficiency to determine inclusion.

Three studies used standardized reading tests to identify struggling readers (Chan, 1991; Nolan, 1991; Wong & Jones, 1982). Chan (1991) recruited participants who read two or more years below expected grade level as measured on the St. Lucia Graded Word Reading Test (Andrews, 1973) and the GAP reading comprehension test (Form B3; McLeod, 1977). Nolan (1991) included students who scored 0.6 to 3.9 years below grade-level on the Gates-MacGinitie Reading Test (Gates & MacGinitie, 1971). Wong and Jones (1982) used the Nelson Reading Skills test (Hanna, Schell, & Schreiner, 1977) and selected students with LD who read three to four years below grade-level while also measuring student IQ on the Wechsler Intelligence Scale for Children (Wechsler, 1974). Two studies (Chan, 1991; Wong & Jones, 1982) confirmed that participants had no physical, sensory, and/or emotional disabilities.

Four studies included participants already identified with LD or as struggling readers receiving special education or remedial services in their school districts (Berkeley et al., 2011; Chan & Cole, 1986; Clark et al., 1984; Rouse et al., 2014). Of the four, two studies, Chan and Cole (1986) and Clark et al. (1984), noted that participants' IQ was in the normal range. One study (Cohen, 1983) included students based on a single criterion: a researcher developed reading pretest. Students who scored below 85% on this test were included while those who scored 85% or higher were excluded.

Grade-level performance. Of the ten studies in this synthesis, five studies measured the effects of self-questioning strategy intervention on elementary level struggling readers' reading outcomes. Effects of the intervention were mixed for this student population. Based on Cohen's (1988) recommendation to interpret effect size, two studies found medium to large positive effects of the intervention (Cohen, 1983; Rouse et al., 2014), while one study (Holmes, 1985) found no significant differences between treatment and control group students. One study (Chan & Cole, 1986) found positive and negative effects of the intervention for different treatment and comparison strata. Finally, one study (Chan, 1991) did not have a true comparison group.

Similarly, effects of self-questioning strategy intervention on reading outcomes for middle and high school students also varied. Of the five studies, only one found (Wong & Jones, 1982) medium positive effects of self-questioning intervention on students' reading comprehension outcomes. Two studies (Berkeley et al., 2011; Nolan, 1991) found no significant differences between treatment and comparison groups, and one reported (Davey & McBride, 1986) mixed results. Finally, for one study (Clark et al., 1984), it was not possible to conduct visual analysis or establish a functional relationship due to lack of data presented in the manuscript.

Treatment and Comparison.

Treatment. Five studies provided explicit instruction—modeling, guided practice, and independent practice—in using self-questioning strategy while reading texts. One single-case design study (Rouse et al., 2014) reported very large effects of the intervention on comprehension outcome measure. However, it was not possible to conduct visual analysis to establish a functional relationship as only two attempts to demonstrate intervention effects were reported; only two participants were included in the multiple baseline design study.

In contrast, three treatment and comparison studies (Berkeley et al., 2011; Holmes, 1991; Nolan, 1991) reported no significant differences between control and treatment groups. Finally, one study (Chan, 1991) did not have a true comparison/business-as-usual (BAU) group as both groups in the study received self-questioning strategy instruction but were assigned to different instructional conditions.

In the remaining five studies, students were instructed in the use of self-questioning strategy but received no modeling or guided practice. While two studies reported medium to large positive effects (Cohen, 1983; Wong & Jones, 1982) that were significant, two studies reported mixed results (Davey and McBride, 1986; Chan & Cole, 1986). In Davey and McBride's study (1986), large positive effects were reported for student performance on an inferential question test measure, however treatment group did not vary significantly from control group on a literal comprehension question posttest measure. Likewise, Chan and Cole (1986) reported significant and non-significant differences between different treatment and control strata on the posttest and transfer measure of reading comprehension. For one study (Clark et al., 1984), no data were available to compute a quantitative measure of difference between participants' baseline and intervention phase results; authors only reported total mean percentage scores for each phase. In summary, difference between treatment and control groups was measured in eight studies. Three studies reported positive significant effects, three studies reported no significant differences between groups, and two studies reported mixed results on the effects of self-questioning strategy instruction on struggling readers' reading outcomes.

Comparison. In Cohen's (1983) study, comparison group students continued to participate in general education classroom instruction along with typically developing peers. Other researchers directed comparison group students to reread text (Chan & Cole, 1986; Dave

and McBride, 1986), read and memorize text (Berkeley et al., 2011), and provide justification for selecting answers post reading and testing (Holmes, 1985). None of the interventions involved feedback to students. Nolan (1991) provided comparison group participants with instruction on ways to identify unknown vocabulary words in texts and taught strategies, in the form of activities, to learn their meaning. Similarly, Wong and Jones (1982) provided decoding and vocabulary support to students while reading texts. After reading the text, students were expected to comment on the quality of the writing. Finally, Chan (1991) did not have a comparison group as both groups received different types of self-questioning instruction.

Design information. Of the ten studies included in the synthesis, seven studies used treatment and comparison group design, two studies were single-case design, and one study employed between group design with no true comparison group. Comparisons on various coded study characteristics are reported.

Dosage. Intervention dosage was reported for nine studies and ranged from 1 to 13.5 total hours of intervention. Six studies also reported instructional group sizes ranging from one-on-one sessions to groups comprising up to five students. Frequency of sessions was reported in seven studies ranging from one per day to one per week. Finally, total number of sessions was reported in eight studies ranging from 3 to 27 sessions.

Attrition. All participants were randomly assigned to treatment or comparison conditions in group design studies. No differential or overall attrition was reported that exceeded the acceptable level (WWC, 2017, p. 11) in all studies included in the synthesis. Group sizes remained similar at the start of the study and during posttest.

Dependent measures. A majority of the studies ($n = 7$) included in this synthesis collected data using researcher-developed reading comprehension measures (Berkeley et al.,

2011; Chan, 1991; Chan & Cole, 1986; Clark et al., 1984; Davey & McBride, 1986; Rouse et al., 2014; Wong & Jones, 1982). Two of the studies (Berkeley et al., 2011; Davey & McBride, 1986) reported test reliability data for researcher developed measures. In addition to researcher developed reading comprehension posttest, Holmes (1991) also administered the Nelson Reading standardized reading comprehension test (Hanna, Schell, & Schreiner, 1977). In contrast, only two studies administered a standardized achievement posttest. Nolan (1991) tested intervention effects using the Stanford Diagnostic test (Karlsen, Madden, & Gardner, 1986), and Cohen (1983) used the general comprehension section of the Developmental Reading test (Bond, Clymer, & Hoyt, 1955).

Research Question 1b: Methodological Rigor of Studies Included

All group design studies included in this synthesis *met standards without reservation*. All group design studies determined group membership through a random process. Additionally, overall and differential attrition was low. In contrast, both single-case design studies *did not meet design standards* due to the lack of three attempts to demonstrate intervention effects.

<div align="center">

Discussion

</div>

The purpose of this study was to analyze the effects of self-questioning strategy interventions on reading comprehension outcomes for students identified as LD or struggling readers in grades K-12. Our literature search of the last 53 years of research yielded ten studies that fit our inclusion criteria. For the eight group-design studies, effect size and statistical significance were calculated to determine the effectiveness of self-questioning strategy intervention. The remaining two studies used single-case designs, one of which provided data to calculate Tau-U effect size for each participant in the study.

The mean weighted effect size for all self-questioning strategy instruction studies on comprehension outcomes for struggling readers was moderately positive. Of the eight group-design studies included, seven studies compared self-questioning strategy instruction to a comparison group. Of these, two studies reported statistically significant effects of self-questioning intervention on students' reading outcomes (Cohen, 1983; Wong & Jones, 1982), three studies reported nonsignificant differences between comparison and treatment groups (Berkeley et al., 2011; Holmes, 1985; Nolan, 1991), and two studies reported mixed results (Chan & Cole, 1986; Davey & McBride, 1986). Additionally, in the one study with no comparison group (Chan, 1991), both groups receiving self-questioning strategy instruction (explicit and non-explicit) did not differ significantly on the posttest reading comprehension measure.

It may be beneficial to note that statistical significance may not have been attained in group design studies due to the relatively small total sample in five of the eight studies (n<30). In three studies that used large (total) samples (n>30), two studies (Cohen, 1983; Wong & Jones, 1982) reported statistically significant positive effects of self-questioning strategy instruction on reading comprehension outcomes. The third study, on the other hand, (Davey & McBride, 1986), reported significant improvement on post-outcome measure of inferential questions but no significant difference on the comprehension measure of literal questions.

Nolan (1991) reported outcome data for different student reading ability levels. While the study reported positive effects of self-questioning intervention on struggling readers' reading outcomes, effects varied for students with differing reading abilities. Large effects were observed for students who read 1.6 to 2 years below grade-level and moderate effects for students who read 0.6 to 1.5 years or 2.2 to 2.9 years below grade-level. However, students who read 3.1 to 3.9

years below grade-level showed small negative effects of the intervention. These results could

imply that self-questioning strategy may be more effective for students who are moderately

below grade-level in reading. Indeed, it could also indicate that students who read three or more

years below grade-level may need more intensive interventions such as increased frequency and

duration of sessions to gain proficiency in strategy-use (Vaughn, Wanzek, Murray, & Roberts,

2012).

Implications for Practice

Self-questioning strategy is a highly recommended reading comprehension strategy in the

education research literature (National Reading Panel, 2000; Willingham, 2006-2007). However,

evidence of the effects of self-questioning strategy instruction for students with LD and

struggling readers has been inconclusive. Results of this review indicates overall positive effects

of self-questioning strategy instruction on reading comprehension outcomes for this population.

However, these results should be interpreted with caution, as the corpus of studies was small

with few participants and many of the studies were not conducted in past ten years.

In the current synthesis, none of the effects of self-questioning strategy intervention were

associated with participants' grade-level and type of instruction. Medium to large effects were

observed for elementary school students' reading outcomes (Chan & Cole, 1986; Cohen, 1982;

Holmes, 1985) and small to large effects for middle to high school students' reading outcomes

(Berkeley et al., 2014; Davey & McBride, 1986; Nolan, 1991; Wong & Jones, 1982). This

finding is aligned with previous research, showing larger effects for reading interventions in the

elementary grades, and smaller effects for studies with students in the secondary grades

(Scammacca et al., 2015; Wanzek et al., 2007).

Small to large effects were reported when teachers at least modeled and provided guided practice in the use of self-questioning strategy (Berkeley et al., 2014; Chan, 1991; Holmes, 1985; Nolan, 1991), and similar effects were observed when instructors only taught students how to generate and answer questions while reading (Cohen, 1983; Davey & McBride, 1986; Wong & Jones, 1982). However, one study reported statistically significant positive effects of explicit instruction on transfer measures when students were not prompted to use the strategy post-intervention (Chan, 1991). While both explicit and non-explicit strategy instruction may be beneficial for improving struggling readers' reading comprehension, explicit strategy instruction may improve generalization and allow students to use self-questioning strategy independently.

Effects of the total number of hours of instruction on students' reading outcomes varied slightly. One study in which students received less than two hours of total strategy-use instruction had small effects (Berkeley et al., 2014). In contrast, medium to large effects were observed for students receiving two or more hours of strategy instruction (Chan & Cole, 1986; Cohen, 1983; Davey & McBride, 1986; Holmes 1985; Nolan, 1991; Wong & Jones, 1982). Our results are consistent with education research literature indicating that longer exposure to interventions have positive effects on students' academic achievements (Denton, Fletcher, Anthony, & Francis, 2006; Vaughn, Linan-Thompson, & Hickman, 2003; Wanzek & Vaughn, 2007).

We suggest that practitioners consider the local context when implementing self-questioning interventions. It is important to determine the amount of time available for intervention, the implementor, age/grade of the participants, and setting. While self-questioning strategy may benefit some students, we recommend that teachers monitor students' comprehension outcomes and if the strategy is not having the desired effect, to consider

alternative reading comprehension strategies. Indeed, teachers may also combine the self-questioning strategy with other reading comprehension strategies (e.g., Hagaman et al., 2016) and evaluate the effectiveness for individual students. Past systematic reviews for this student population have shown that combining self-questioning strategy with paragraph restatement / summarization (Sencibaugh, 2005), main idea generation, and text structure analysis (Berkeley, Scruggs, & Mastropieri, 2010) have yielded positive outcomes.

Limitations

This synthesis has a few limitations that need to be acknowledged. One key limitation in interpreting the results of this study is that only ten studies met criteria to be included in the synthesis. The paucity of studies found could be attributed to difficulty in publishing studies that report nonsignificant results, also known as the "file drawer" problem (Rosenthal, 1979). One more reason for a small corpus of studies could be that more reading intervention studies are designed with multicomponent reading strategies (Scammacca et al., 2016) that make it difficult to isolate and analyze single components, and lead to fewer studies being identified for inclusion. Another common theme in educational intervention studies is the dearth of well-designed research interventions (Cook, 2007). However, it should be noted that a majority of studies included in this synthesis met all of the WWC criteria for study design. A further limitation was the small total sample size ($n < 30$) in five of the eight group-design studies. Small total sample size may have decreased the chance of detecting significant effects and may also have inhibited the ability to generalize the findings.

Conclusion and Future Research

This synthesis did not find conclusive evidence of the effectiveness of self-questioning strategy use in improving reading comprehension outcomes for students with LD and struggling

readers in grades K-12. While self-questioning strategy instruction has been widely

recommended (e.g., National Reading Panel, 2000; Willingham, 2006-2007), its effectiveness for

struggling readers remains a phenomenon that warrants further investigation. Recent studies of

reading interventions for this population have examined strategy packages that incorporate

several strategies without looking at the individual effects of each strategy. One direction for

future research is to examine the individual components of these interventions and to test their

efficacy in isolation and in combination with each other in order to determine the most effective

components for struggling readers and students with LD. It would be valuable for future

researchers to investigate the following questions with large sample sizes to test the efficacy of

self-questioning strategy use for improving students with LD or struggling readers'

comprehension outcomes:

1. Does the effectiveness of self-questioning vary based on text difficulty (grade-level vs.
   independent-level) and/or text content (narrative vs. expository)?

2. How effectively can students sustain self-questioning strategy use post intervention to
   enhance their reading comprehension outcomes?

3. Are the effects of self-questioning strategy use similar or different for students in
   elementary, middle, and high school?

4. Does the length of the intervention moderate the effects of self-questioning strategy
   intervention on reading comprehension outcomes for students with LD and/or struggling
   readers?

References

References marked with an asterisk indicate studies included in this synthesis

Andre, M. E., & Anderson, T. H. (1978). The development and evaluation of a self-questioning study technique. *Reading Research Quarterly*, 605-623.

Andrews, R. J., & Andrews, R. J. (1973). St. Lucia graded word reading test. Teaching and Testing Resources.

Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. *Handbook of Reading Research*, Volume 1 (353), V394.

*Berkeley, S., Marshak, L., Mastropieri, M. A., & Scruggs, T. E. (2011). Improving student comprehension of social studies text: A self-questioning strategy for inclusive middle school classes. *Remedial and Special Education*, *32*(2), 105-113.

Berkeley, S., Scruggs, T. E., & Mastropieri, M. A. (2010). Reading comprehension instruction for students with learning disabilities, 1995—2006: A meta-analysis. *Remedial and Special Education*, *31*(6), 423-436. doi:10.1177/0741932509355988

Bernstein, S. (1973). The effects of children's question-asking behaviours on problem solution and comprehension of written materials. Unpublished doctoral dissertation, Columbia University.

Bond, G. L., Clymer, T., & Hoyt, C. J. (1955). The Developmental Reading Tests. Chicago: Lyons and Carnahan. Chicago, IL.

California Test Bureau. California Achievement Tests. Monterey, California: CTBI McGraw-Hill, 1978.

*Chan, L. K. S. (1991). Promoting strategy generalization through self-instructional training in

    students with reading disabilities. *Journal of Learning Disabilities*, *24*(7), 427-433.

    doi:10.1177/002221949102400708

*Chan, L. K. S., & Cole, P. G. (1986). The effects of comprehension monitoring training on the

    reading competence of learning disabled and regular class students. *Remedial and Special*

    Education, *7*(4), 33-40. doi:10.1177/074193258600700407

*Clark, F. L., Deshler, D. D., Schumaker, J. B., Alley, G. R., & Warner, M. M. (1984). Visual

    imagery and self-questioning: Strategies to improve comprehension of written material.

    Journal of Learning Disabilities, *17*(3), 145-149. doi:10.1177/002221948401700304

Clearinghouse, W. W. (2017). *Procedures and standards handbook (Version 4.0).* Washington,

    DC: Institute of Education Sciences.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ:

    Erlbaum.

*Cohen, R. (1983). Self-generated questions as an aid to reading comprehension. *The Reading*

    Teacher, *36*(8), 770-775.

Collins, A., Brown, J.S., & Larkin, K.M. (1980). Inference in text understanding. *Theoretical*

    Issues in Reading Comprehension, 385-407.

Cook, T. D. (2007). Randomized experiments in education: Assessing the objections to doing

    them. *Economics of Innovation and New Technology*, *16*(5), 331-355.

Crabtree, T., Alber-Morgan, S. R., & Konrad, M. (2010). The effects of self-monitoring of story

    elements on the reading comprehension of high school seniors with learning disabilities.

    Education and Treatment of Children, *33*(2), 187-203.

*Davey, B., & McBride, S. (1986). Generating self-questions after reading: A comprehension assist for elementary students. *The Journal of Educational Research*, *80*(1), 43-46.

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.

Duffy, G. G., & Roehler, L. R. (1987). Improving reading instruction through the use of responsive elaboration. *The Reading Teacher*, *40*(6), 514-520.

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*(1), 262-300.

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*(1), 37.

Frase, L. T., & Schwartz, B. J. (1975). Effect of question production and answering on prose recall. *Journal of Educational Psychology*, *67*(5), 628.

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Martinez, E. A. (2002). Preliminary evidence on the social standing of students with learning disabilities in PALS and no–PALS classrooms. *Learning Disabilities Research & Practice*, *17*(4), 205-215.

Gast, D. L., & Ledford, J. R. (Eds.). (2014). *Single case research methodology: Applications in special education and behavioral sciences.* New York, NY: Routledge.

Gates, A. I., & MacGinitie, W. H. (1971). *Gates-MacGinitie Reading Tests.* Teachers College Press, Columbia University.

Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research*, *71*(2), 279-320. doi:10.3102/00346543071002279

Graesser, A. (2012). An introduction to strategic reading comprehension. In D.S. McNamara (Ed.), *Reading Comprehension Strategies: Theories, Interventions, and Technologies* (3-26). NJ: Lawrence.

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*(1), 43-52. doi:10.1080/1045988X.2014.966802

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

*Holmes, B. C. (1985). The effects of a strategy and sequenced materials on the inferential-comprehension of disabled readers. *Journal of Learning Disabilities*, *18*(9), 542-546. doi:10.1177/002221948501800909

Hanna, G., Schell, L. M., & Schreiner, R. (1977). The Nelson reading skills test. Itasaca, IL: Riverside.

Karlsen, B., Madden, R., & Gardner, E. F. (1986). Stanford Diagnostic Reading Test: Brown Level Forms GH. Psychological Corporation.Chicago

Klingner, J. K., & Vaughn, S. (1998). Using collaborative strategic reading. *Teaching Exceptional Children*, *30*(6), 32-37.

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 What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Sage Publications, Inc.

Mastropieri, M. A., & Scruggs, T. E. (1997). Best practices in promoting reading comprehension
in students with learning disabilities 1976 to 1996. *Journal for Special Educators*, *18*(4),
198-213.

McCardle, P., Scarborough, H. S., & Catts, H. W. (2001). Predicting, explaining, and preventing
children's reading difficulties. *Learning Disabilities Research & Practice*, *16*(4), 230-
239.

McLeod, J. (1977). GAP reading comprehension test (3rd ed.). Melbourne: Heinemann
Educational Australia.

Nation, K. (2005). Children's reading comprehension difficulties. In M. Snowling & C. Hulme
(Eds.) *The science of reading: A handbook* (pp. 248–266). Boston: Blackwell Synergy.

National Reading Panel (2000). *Teaching children to read: An evidence-based
assessment of the scientific research literature on reading and its implications for
reading instruction.* Washington, D.C.: National Institute of Child Health
and Human Development.

*Nolan, T. E. (1991). Self-questioning and prediction: Combining metacognitive strategies.
Journal of Reading, *35*(2), 132-138.

Ogle, D. M. (1986). KWL: A teaching model that develops active reading of expository text. *The
Reading Teacher, 39*(6), 564-570.

Paris, S. G., Wasik, B. A., Turner, J. C., (1991). The Development of Strategic Readers. R. T.
O'Connell (Ed), *Handbook of Reading Research, Volume II* (pp. 609-640). White Plains,
NY: Longman.

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and

    trend for single-case research: Tau-U. *Behavior Therapy*, *42*(2), 284-299.

Pearson, D., & Fielding, L. (1991). Comprehension Instruction. In R. Barr et al. (Eds.) *Handbook*

    of reading research volume 2 (pp. 815–850). NY: Longman.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively*

    responsive reading. New York: Routledge.

Pressley, M., Borkwski, J. G., & Schneider, W. (1989). Good information processing: What it is

    and how education can promote it. *International Journal of Educational Research*, *13*(8),

    857-867.

Pustejovsky, J. E. (2017). Single-case effect size calculator (Version 0.2) Web application.

    Retrieved from https://jepusto.shinyapps.io/SCD-effect-sizes/

    (https://jepusto.shinyapps.io/SCD-effect-sizes/)

Robinson, F.P. (1941) *Diagnostic and remedial techniques for effective study.* New York, NY:

    Harper & Brothers.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A

    review of the intervention studies. *Review of Educational Research*, *66*(2), 181-221.

    doi:10.2307/1170607

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological*

    Bulletin, *86*(3), 638-641. doi:10.1037//0033-2909.86.3.638

*Rouse, C. A., Alber-Morgan, S. R., Cullen, J. M., & Sawyer, M. (2014). Using prompt fading to

    teach self-questioning to fifth graders with LD: Effects on reading comprehension.

    Learning Disabilities Research & Practice, *29*(3), 117-125.

Scammacca, N. K., Roberts, G. J., Cho, E., Williams, K. J., Roberts, G., Vaughn, S. R., &
Carroll, M. (2016). A century of progress: Reading interventions for students in grades 4–
12, 1914–2014. *Review of Educational Research*, *86*(3), 756-800.

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*(4), 369-390.

Schumaker, J. B., Denton, P. H., & Deshler, D. D. (1984). *The paraphrasing strategy*. Lawrence:
University of Kansas.

Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C., &
Torgesen, J. (2010). *Improving reading comprehension in kindergarten through 3rd
grade: A practice guide* (NCEE 2010-4038). Washington, DC: National Center for
Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S.
Department of Education.

Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P.,
... & Fletcher, J. M. (2004). Development of left occipitotemporal systems for skilled
reading in children after a phonologically-based intervention. *Biological Psychiatry*,
55(9), 926-933.

Simpson, P. S. (1989). The effects of direct training in active comprehension on reading
achievement, self-concepts, and reading attitudes of at-risk sixth grade students.
Unpublished doctoral dissertation, Texas Tech University.

Smith, A. E. (1972). The effectiveness of training students to generate their own questions prior
to reading. Unpublished doctoral dissertation, Syracuse University.

Smith, N. J. (1977). The effects of training teachers to teach students at different reading ability

      levels to formulate three types of questions on reading comprehension and question

      generation ability. Unpublished doctoral dissertation, University of Georgia

Snow, C. (2002). Reading for understanding: Towards a R&D program in reading

      comprehension. Washington, DC: RAND Reading Study Group.

Taylor, L. K., Alber, S. R., & Walker, D. W. (2002). The comparative effects of a modified self-

      questioning strategy and story mapping on the reading comprehension of elementary

      students with learning disabilities. *Journal of Behavioral Education*, *11*(2), 69-87.

Thorndike, E. L. (1917). Reading as reasoning: A study of mistakes in paragraph reading.

      Journal of Educational Psychology, *8*(6), 323.

Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K., & Conway,

      T. (2001). Intensive remedial instruction for children with severe reading disabilities:

      Immediate and long-term outcomes from two instructional approaches. *Journal of*

      Learning Disabilities, *34*(1), 33-58.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., &

      Garvan, C. (1999). Preventing reading failure in young children with phonological

      processing disabilities: Group and individual responses to instruction. *Journal of*

      Educational Psychology, *91*(4), 579.

Vannest, K. J., & Ninci, J. (2015). Evaluating intervention effects in single-case research

      designs. *Journal of Counseling & Development*, *93*(4), 403-411.

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: The Meadows

      Center for Preventing Educational Risk

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.

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

Wanzek, J., & Vaughn, S. (2007). Research-based implications from extensive early reading interventions. *School Psychology Review*, *36*(4), 541.

Wagoner, S. A. (1983). Comprehension monitoring: What it is and what we know about it. *Reading Research Quarterly*, *23*, 328–346.

Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children–Revised.* New York, NY: Psychological Corporation.

Wilkinson, I. A., & Son, E. H. (2011). A Dialogic Turn in Research on Learning and Teaching to Comprehend. In M. L. Kamil et al. (Eds.), *Handbook of reading research volume 4* (pp. 359-387). New York, NY: Routledge.

Willingham, D. T. (2006-2007). The usefulness of brief instruction in reading comprehension strategies. *American Educator, 30*(4), 39-50.

Wong, B. Y. (1985). Self-questioning instructional research: A review. *Review of Educational Research*, *55*(2), 227-268. doi:10.2307/1170191

*Wong, B. Y., & Jones, W. (1982). Increasing metacomprehension in learning disabled and normally achieving students through self-questioning training. *Learning Disability Quarterly*, *5*(3), 228-240.

Table 1. Study information.

Study ▪ Participant disability ▪ Participant grade	Design Information ▪ Design ▪ Dosage (total hours per student) ▪ Frequency/Total Sessions	Treatment (*n*=sample size)	Comparison (*n*=sample size)
Berkeley et al., 2011 ▪ Learning disability ▪ 7th graders	▪ Experimental (Treatment/Comparison) ▪ 1 h ▪ 1 per day / 3	Explicit instruction in how to apply SQ strategy while reading text (SQ; *n* = 2)	ELA instruction as usual, students asked to read and memorize as much as possible (CO; *n* = 3)
Cohen, 1983 ▪ Struggling readers ▪ 3rd graders	▪ Experimental (Treatment/Comparison) ▪ 2.83 h ▪ 1 per day / 11	Questions generation instruction booklet with examples followed by instruction on question generation while reading text (SQ; *n* = 23)	ELA instruction as usual (CO; *n* = 24)
Chan, 1991 ▪ Reading disability ▪ 5th and 6th graders	▪ Experimental (Between group design) ▪ 3.33 h ▪ 1 per day / 5	Explicit instruction in how to apply SQ strategy while reading text (SQ1; *n* = 10) [1]Cued condition: n=5 [1]Uncued condition: n=5	Teacher demonstrated the use of SQ strategy followed by independent practice (SQ2; *n* = 10) [1]Cued condition: n=5 [1]Uncued condition: n=5
Chan and Cole, 1986 ▪ Learning disability ▪ 5th and 6th graders	▪ Experimental (Treatment/Comparison) ▪ 2 h ▪ 1 per day / 4	Students were taught to generate questions while reading a text. They were to think of at least two questions. School 1 (SQ1; $n = 4^2$) School 2 (SQ2; $n = 4^2$)	Students read and then re-read each paragraph to the robot. School 3 (C01; $n = 4^2$) School 4 (C02; $n = 4^2$)

Study ▪ Participant disability ▪ Participant grade	Design Information ▪ Design ▪ Dosage (total hours per student) ▪ Frequency/Total Sessions	Treatment (n=sample size)	Comparison (n=sample size)
Davey and McBride, 1986 ▪ Below-average, average, and above-average readers	▪ Experimental (Treatment/Comparison) ▪ 6th graders ▪ NR ▪ NR	Instruction to generate two questions that assessed important ideas of the passage. (SQ; $n = 24$)	Students were asked to read, reread, and study the text thoroughly (CO; $n = 26$)
Holmes, 1985 ▪ Struggling readers ▪ 4th and 5th graders	▪ Experimental (Treatment/Comparison) ▪ 2.67 h ▪ NR / 8	Instructors modeled and provided guided practice on how to use SQ strategy (SQ; $n = 6$)	ELA instruction as usual (CO; $n = 6$)
Nolan, 1991 ▪ Struggling readers ▪ 6th, 7th, and 8th graders	▪ Experimental (Treatment/Comparison) ▪ 3 h ▪ 1 per week / 3	Explicit instruction in how to apply SQ strategy while reading text (SQ; $n = 14$)	Identifying unknown vocabulary words while reading and learning their meaning (CO; $n = 14$)
Wong and Jones, 1982 ▪ Learning disability ▪ 8th and 9th graders	▪ Experimental (Treatment/Comparison) ▪ 7 h ▪ One per day / 5	Instruction to find main idea, create questions about main idea, and answer self-created questions (SQ; $n = 30$)	ELA instruction as usual (CO; $n = 30$)
Rouse et al., 2014 ▪ Learning disability ▪ 5th graders	▪ Single case design (Multiple baseline across participants) ▪ 11.5 to 13.5 h ▪ NR / 23 to 27	Explicit and structured instruction in how to apply SQ strategy while reading text (SQ; $n = 2$)	N/a

Study
▪ Participant disability
▪ Participant grade
	Design Information
	  ▪ Design
	  ▪ Dosage (total hours per student)
	  ▪ Frequency/Total Sessions
		Treatment
		(*n*=sample size)
			Comparison
			(*n*=sample size)

Clark et al., 1984 ▪ Learning disability ▪ 8th, 9th, and 11th graders	▪ Single case design (Multiple baseline across two strategies) ▪ 5 to 7 h ▪ One per day to one per week / NR	Instruction on use of Wh questions and question generation (SQ; *n* = 6)	N/a

NR=Not reported; N/a = Not applicable; SQ = Self-questioning strategy; CO = Comparison group.

1. Cued condition: Students were prompted to use the strategy learnt before taking the test; Uncued condition: Students were not prompted to use the strategy learnt before taking the test.

2. Reported sample size for each group ranged from four to five. To calculate an estimated effect size, sample size was assumed to be four participants in each group.

Table 2. Study measures and outcomes.

Study	Dependent Measure(s) (Reading Comprehension test)	Std	Results	Effect Size - 95% Confidence Interval
Berkeley et al., 2011	Posttest 1, multiple choice questions	N	SQ < CO: g = -0.20	-1.99 to 1.60
	Posttest 2, open-ended questions	N	SQ > CO: g = 0.09	-1.70 to 1.88
Cohen, 1983	Developmental Reading Test, General Comprehension Section	Y	SQ > CO: g = 0.71[*]	0.12 to 1.30
Chan, 1991	Posttest 1, multiple-choice test	N	SQ1>SQ2: g = 0.20	-1.12 to 1.53
	Posttest 2a, main idea test – Cued condition	N	SQ1>SQ2: g = 0.26	-0.98 to 1.51
	Posttest 2b, main idea test – Uncued condition	N	SQ1>SQ2: g = 1.81[*]	0.33 to 3.28
Chan and Cole, 1986	Posttest, multiple-choice test	N	SQ1>CO1: g = 1.34	-0.19 to 2.88
			SQ1>C02: g = 1.72[*]	0.10 to 3.35
			SQ2>CO1: g = 1.52	-0.06 to 3.09
			SQ2>CO2: g = 2.01[*]	0.31 to 3.71
	Transfer measure, multiple-choice test	N	SQ1>CO1: g = 0.06	-1.33 to 1.45
			SQ1<CO2: g = -0.39	-1.79 to 1.00
			SQ2>CO1: g = 3.77[*]	1.46 to 6.08
			SQ2>CO2: g = 1.67[*]	0.06 to 3.28
Davey and McBride, 1986	Posttest 1, literal questions	N	SQ < CO: g = -0.06	-0.62 to 0.49
	Posttest 2, inferential questions	N	SQ > CO: g = 1.08[*]	0.49 to 1.67
Holmes, 1985	Posttest, inference questions	N	SQ > CO: g = 0.30	-0.84 to 1.44
	Nelson Reading Test	Y	SQ > CO: g = 0.74	-0.43 to 1.91
Nolan, 1991	Stanford Diagnostic Reading Test	Y	SQ > CO: g = 0.45	-0.30 to 1.20

Study	Dependent Measure(s) (Reading Comprehension test)	Std	Results	Effect Size - 95% Confidence Interval
Wong and Jones, 1982	Posttest, open ended questions	N	SQ > CO: $F(1, 56) = 4.71, p<.05$	N/a
Rouse et al., 2014	Posttest, multiple choice questions	N	Tau-U Student 1 = 1.00[1] Student 2 = 1.02[1]	N/a
Clark et al., 1984	Text used: Pretest 1, Ability level[3] Posttest 1, Ability level[3] Pretest 2, Grade level[3] Posttest 2, Grade level[3] Maintenance, Ability[3] Maintenance, Grade[3]	N	Mean Correct[2] 52.7% 89.6% 46.0% 89.8% 72.0% 70.0%	N/a

N/a = Not applicable; SQ = Self-questioning strategy; CO = Comparison group. Std = Standardized Measure.

*$p<0.05$.

1. Tau-U effect size measure shows the difference between students' baseline phase data and final self-questioning fading phase data.

2. Researchers only reported aggregate student results for each phase and type of text used. Individual student data were not available for analysis.

3. Ability level is synonymous with students' independent reading level text, and grade level is synonymous with students' grade level text.

Appendix A

Interobserver agreement by Response Type

Response Type	Pre-discussion IOA (%)	Post-discussion IOA (%)
Forced Choice	96.71	100
Open-ended	95.26	100
Numeric Data Entry	96.48	100
Total	96.15	100

Examples: Forced choice: Yes/No; Open-ended: Describe intervention; Numeric Data Entry: Mean, Standard deviation, sample size.