Critical Factors in Reading Comprehension Instruction for Students with Learning Disabilities: A Research Synthesis

Woori Kim
Seoul National University

Sylvia Linan-Thompson
The University of Texas at Austin

Radhika Misquitta
The Gateway School of Mumbai

This review examined the effectiveness of critical factors in instruction for improving the reading comprehension of middle school students with learning disabilities. Five critical factors were identified: (i) type of instructional methods, (ii) self-monitoring, (iii) components of reading incorporated, (iv) fidelity of instruction (scripted vs. nonscripted and researcher vs. teacher), and (v) group size. Fourteen studies published between 1990 and 2010 were reviewed. The findings indicated that interventions incorporating strategy instruction, specifically, main idea and summarization, yielded high effects on comprehension. The use of self-monitoring combined with main idea strategy improved comprehension performance. It was found that both, instruction targeting comprehension alone, as well as comprehension that incorporated other reading components such as vocabulary had significant effects on comprehension. Researcher-delivered instruction with script yielded higher effects across intervention types. One-on-one instruction or instruction in pairs was more effective than large group instruction across intervention types.

Reading comprehension difficulties for students with learning disabilities (LD) have repeatedly been documented in the literature. As students with LD move from elementary to middle school grades, these difficulties become more acute, as students are expected to “read to learn” instead of “learn to read.” Identifying “what works” in teaching reading comprehension is an essential area of reading research, especially as students move to higher academic grades.

Several strategies have been identified to improve comprehension for students with LD, such as the use of a self-monitoring strategy, identifying the main idea, using inferences, the use of semantic mapping and graphic organizers (GOs), and reciprocal teaching, to name a few (Gajria, Jitendra, Sood, & Sacks, 2007). However, despite advances in our understanding of effective practices to improve reading comprehension, the National Assessment of Educational Progress (NAEP) reported that problems of lower performing students in reading comprehension persist, with only a slight percentile gain seen from 2002 to 2009 for such students (from 244 to 243 for those at 25th percentile and from 220 to 219 for those at 10th percentile).

One possible explanation for this poor performance could be that while several studies have examined the effectiveness of instructional methods to enhance reading comprehension for elementary students with disabilities, much less is known about effective instruction specifically for middle- and high school students (Curits & Longo, 1999). Previous syntheses on effective reading comprehension instruction have tended to collapse data for students of all ages (see Gersten, Fuchs, Williams, & Baker, 2001; Gajria et al., 2007; Swanson & Hoskyn, 1998). This practice has several shortcomings, as strategies effective for younger readers may not be applicable to older students, given the developmental differences as well as the demands of reading tasks (Edmonds et al., 2009).

Edmonds et al. (2009) have attempted to fill this gap in the literature by examining interventions specifically for older struggling readers. In their meta-analysis, Edmonds et al. identified strategies to enhance reading comprehension such as explicit instruction in reading comprehension, employing multiple strategies, and emphasizing comprehension versus word attack skills or fluency. Our article seeks to further tease out components of these interventions that may be effective in improving reading comprehension, as well as to narrow the scope of results to only students with LD. We selected five intervention components documented in the literature (Fuchs & Fuchs, 2007; Wanzek & Vaughn, 2007) as critical factors based on a review of relevant literature: (i) type of instructional methods, (ii) self-monitoring, (iii) components of reading incorporated, (iv) fidelity of instruction (scripted vs. nonscripted and researcher vs. teacher), and (v) group size.

Requests for reprints should be sent to Woori Kim, The University of Texas at Austin. Electronic mail may be sent to kworim@utexas.edu.
The Five Critical Factors in Comprehension Instruction

Type of Instructional Methods

Several reviews have identified instructional methods that improved reading comprehension for students with LD (Gajria et al., 2007; Gersten, Fuchs, Williams, & Baker, 2001; Kim, Vaughn, Wanzek, & Wei, 2004). Understanding which instructional methods to employ at a given point is important to help teachers respond appropriately to student needs.

Instructional methods may fall under two categories: instructional modifications, or strategy instruction (Misquita, 2009). Instructional modifications are those interventions that modify the content for the student to make it more accessible. Strategy instruction on the other hand, provides students with the tools to decode the text themselves. An example of instructional modifications is making use of GOs, or supplemental reading and video materials. Strategy instruction may include teaching students to identify the main idea, questioning, and summarization. Studies, incorporating both instructional modifications and strategy instruction, have demonstrated high effect sizes to improve reading comprehension for students with LD (Gajria et al., 2007). This review hopes to guide teachers on how to select among various instruction methods, by identifying what components can be varied and how.

Self-Monitoring

Reviews of the literature have identified self-monitoring as associated with improvement in reading comprehension (Mastropieri & Scruggs, 1997; Reid, 1996). Self-monitoring involves the active participation of students in learning and continuous engagement in activities (Torgesen, 1982). Research on reading comprehension recommends the use of self-monitoring strategy in combination with other strategies (Gersten et al., 2001). For example, Graves (1986) compared the effects of direct instruction versus direct instruction plus self-monitoring to identify the main idea of a passage. The findings indicated that direct instruction plus self-monitoring strategy was more effective than a direct instruction-only strategy. While, in general, self-monitoring in conjunction with other reading strategies is seen to be more effective, results differ depending on outcome measures employed (Jitendra, Cole, Hoppes, & Wilson, 1998; Malone & Mastropieri, 1991; Mastropieri & Scruggs, 1997). These equivocal findings require a more a structured review to better understand the impact of self-monitoring on comprehension.

Reading Components included in Intervention

The National Reading Panel (NRP) (U.S. Department of Health and Human Services, National Institutes of Health, 2000) identified five essential components of reading: phonemic awareness, phonics, fluency, vocabulary, and comprehension. To achieve reading comprehension, which is the highest level of the reading components, students need to be proficient in lower levels of reading such as word decoding and fluency (RAND Reading Study Group, 2002). Researchers, however, differ in their opinions of the optimal combination of reading components to improve reading comprehension (Edmonds et al., 2009). One group calls for focused instruction where instruction in reading comprehension targets only comprehension, and does not include other components of reading. Researchers (Buly & Valencia, 2003; Pany & Jenkins, 1978; Pany, Jenkins, & Schreck, 1982) argue that problems in reading comprehension are the result of a lack of development in fluency and comprehension, and that lower level skills such as the mastery of phonics or word reading would not influence comprehension achievement. Studies on the effects of word recognition are not conclusive on the impact of word recognition skills on reading comprehension (see Fleishler & Jenkins, 1983; Jenkins, Larson, & Fleishler, 1983).

Another body of researchers calls for more balanced instruction, incorporating other components of reading to enhance comprehension. These researchers argue that reading components are correlated and develop together, especially for adolescents with reading problems (Hock et al., 2009), and therefore should be incorporated as a package. Understanding whether or not to include other components of reading, and to what extent, would be beneficial to teachers when selecting among and delivering reading comprehension strategies.

Fidelity of Instruction

In intervention research, it is critical to understand how well instructional methods reflect the goals of the interventions and whether the interventions are delivered as intended (Mathes et al., 2005) in order to draw appropriate conclusions linking the intervention to its effects. Fidelity of instruction refers to whether or not the instruction was delivered as intended. We identified two factors that may contribute to the fidelity of the instruction: if instruction was scripted or not, and whether it was delivered by a researcher or a teacher. Wanzek and Vaughn (2007) compared the effect of two response-to-intervention (RTI) models: standard protocol model and problem-solving model. The authors included a specifically trained instructor as one of the important factors of the standard protocol model. Another study, Mathes et al. (2005) investigated the relative effects of prosaic reading that made use of scripted instruction and responsive reading, with no scripts. Both studies concluded that scripted instruction was not necessary. However, the participants in those studies were early elementary school students. Also, these studies compared the effects of two different intervention models. This limited information needs further review, extending the boundary of age variables and reading components but controlling for other variations in intervention such as instruction methods.

Group Size

Several grouping formats have been adopted to teach reading to students with LD, for example, one-on-one tutoring...
(Elbaum, Vaughn, & Hughes, 1999; Elbaum, Vaughn, Hughes, Moody, & Schuman, 2000), paired instruction (Moody, Vaughn, & Schuman, 1997), small-group formats (Vaughn et al., 2003), and whole group instruction. Previous studies have examined the relative efficacy of different grouping formats (see Elbaum et al., 2000; Vaughn et al., 2003). In their meta-analysis examining the effectiveness of different grouping formats for students with LD, Elbaum et al. (2000) noted that one-to-one tutoring was effective in improving reading outcomes for children with reading difficulties. However, the authors reported no difference between one-to-one and small group instruction. In a study comparing different grouping formats, Vaughn et al. (2003) also reported no significant difference between small group and one-to-one instruction, but reported significant differences between one-to-one and large group ($n = 10$), in favor of one-to-one, and significant differences between small group versus large group instruction, in favor of the small group. Although previous syntheses have examined the relative efficacy of different grouping formats, students included in the analysis are primarily from elementary and upper elementary grades. It would be beneficial for teachers to understand how grouping impacts reading comprehension instruction in middle school settings.

Rationale

Fisher and Frey (2008) posited that poor comprehension outcomes for students with disabilities could be the result of an emphasis of comprehension instruction research on the development of, rather than the application and adaptation of instructional methods, in a given context. Understanding how to vary strategies would enable teachers to maximize their effects in a given condition (Bulgren, Deshler, Schumaker & Lenz, 2000; Heller & Greenleaf, 2007). The purpose of this article, therefore, is to serve as a guide to practitioners when applying reading comprehension strategies in their middle school classrooms and enable them to differentiate instruction and tailor instruction to students’ needs. Specifically, this review seeks to answer the following question:

- What are the effects of critical factors in intervention on the reading comprehension of middle school students with LD?

METHOD

Data Collection

We conducted our search in two parts: (1) an electronic search of educational databases including Academic Search Alumni Edition, Academic Search Complete, ERIC, and PsycINFO using the key terms of comprehension, LD, and middle school, and (2) examination of articles reviewed in previously published syntheses (see Dexter & Hughes, 2011; Gajria et al., 2007; Kim et al., 2004; Swanson, 2008). Studies were limited to those published from 1990 to 2010 in peer-reviewed journals. This initial search yielded 52 articles.

Next, we reviewed titles and abstracts to identify those articles that met the following inclusion criteria. First, studies were included if their purpose was to improve reading comprehension. Studies whose goal was improvement of other reading areas such as phonology or word decoding were excluded. Second, studies where the participants were middle school students (grades 6–8) classified as LD according to the federal and state criteria for eligibility of LD. Studies that included English language learners were excluded since their unique characteristics such as limited language proficiency may not be generalizable to all children with LD. Third, only studies that employed experimental, or quasi-experimental designs and included a control or comparison group were included. Fourth, studies assessed intervention fidelity. Fifth, studies used standardized- or researcher-developed comprehension measures. Sixth, studies provided sufficient quantitative information (e.g., degree of freedom [df], means [M], and standard deviation [SD]) to calculate unbiased effect size (d using Hedges’ $g$). On the basis of these criteria, 14 studies were selected for further review.

Coding Procedures

Each article was coded for the purpose of the study, design, participants (age or grade, disabilities), intervention variables (types and reading components incorporated), intervention delivery (use of script and grouping), dependent measures, and reported findings including $M$ and $SD$. This information was coded by two of the three authors for all of the studies. Inter-rater agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements. Inter-rater agreement of .97 was found. As for the disagreed codes, we resolved the ambiguity after discussion and reached a decision by consensus.

Effect Size Calculation

We calculated the effect size of each dependent measure using means and standard deviations present in studies. In each measure, we calculated the difference between intervention group and comparison group by subtracting the posttest mean score of the control group from the posttest mean score of the intervention group. The difference was then divided by the pooled (average) standard deviation ($s_p$) of the two groups.

$$s_p^2 = \frac{(n_c - 1)SD_c^2 + (n_i - 1)SD_i^2}{n_c + n_i - 2}$$

where 

- $S_{p}$ is the pooled standard deviation,
- $n_i$ is the number of participants in the intervention group,
- $n_c$ is the number of participants in the control group,
- $SD_i$ is the standard deviation of the posttest score of the intervention group, and
- $SD_c$ is the standard deviation of the posttest score of the control group.
Then, to control for bias due to sample sizes, we used Hedges’ g (Hedges & Olkin, 1985). The unbiased effect size estimate was calculated using the following formula:

\[ d = \left(1 - \frac{3}{4N - 9}\right) \left(\frac{M_E - M_c}{s}\right), \]

where 
- \(d\) is the unbiased effect size,
- \(N = n_e + n_c - 2\),
- \(M_E\) is the mean posttest score of the intervention group,
- \(M_c\) is the mean posttest score of the control group.

Effect sizes provided a standardized score for which the magnitude of difference could be determined. We interpreted effect sizes referring the criteria identified by Hedges and Olkin (1985): (i) .80 or greater is a high effect, (ii) at or near .50 is a medium effect, and (iii) .20 or below is a low effect.

RESULTS

From the search, 14 studies met the requisite criteria for this synthesis. Three of these studies were reported in one article (Bakken, Mastroppietri, & Scruggs, 1997) and another article (Bakken et al., 1995). Of the 14 studies, 12 employed random assignment and another article (Bakken et al., 1995). Across those studies, we identified unbiased effect size (ES) on the comprehension measures and reviewed the ESs by measurement types: researcher-developed comprehension posttest (posttest), researcher-developed comprehension follow-up test (follow-up test), and standardized comprehension test (standardized test). We interpreted the effects using the ESs calculated only with the formula presented above.

We organized results into two major sections: brief features of the studies and findings by analyzing data. The findings were reported by five critical factors in order: (i) type of instructional methods, (ii) self-monitoring, (iii) reading components incorporated, (iv) fidelity of instruction, and (v) group size.

Study Features

The studies included a total of 465 students with LD. Of the 14 studies, 12 employed random assignment and another article (Bakken et al., 1995). Studies 1, 2; Boyle, 1996, 2010; Bos & Anders, 1990, Studies 1, 2, 3; DiCecco & Gleason, 2002; Gersten, Baker, Smith-Johnson, Dimino, & Peterson, 2006; Jitendra, Hoppes, & Xin, 2000; Mastroppietri et al., 2001; Wilder & Williams, 2001) and the other two used matched-paired assignment (Calhoon, 2005; Kim et al., 2006). We examined reading comprehension outcomes of researcher-developed measures and standardized measures. Two studies used standardized measures and researcher-developed comprehension measures together (Boyle, 1996; Kim et al., 2006); one (Calhoon, 2005) used a standardized test only; and the others employed only researcher-developed measures.

Effects by Type of Instructional Methods

This review examined the effects by instructional methods used in each study. We organized studies based on whether they incorporated instructional modifications or strategy instruction.

Instructional Modifications

Six of a total of 14 studies employed instructional modifications (Boyle, 1996; Bos & Anders, 1990 Studies, 1, 2, 3; DiCecco & Gleason, 2002; Gersten et al., 2006). Five of the six studies employed GOs (Boyle, 1996; Bos & Anders, 1990, Studies, 1, 2, 3; DiCecco & Gleason, 2002). GOs in those studies were employed in various ways such as cognitive mapping with a mnemonic, semantic mapping, or teacher- and student-developed GO. All studies administered posttests. Boyle added a standardized test to the posttests and Bos and Anders added follow-up tests. Table 1 provides the ESs on the measures for each of the studies. Overall, the effects of GO on comprehension were inconsistent across the studies and measures.

Boyle (1996) examined the effects of cognitive mapping with a mnemonic on comprehension. The intervention group performed high on posttests that assessed literal-comprehension and inferential-comprehension (ES = 1.33 and 0.96, respectively), but lower on the standardized test (ES = 0.34). Bos and Anders (1990) examined the effects of semantic mapping in improving comprehension and vocabulary abilities, compared to the comparison group receiving vocabulary definition instruction without mapping. They conducted three studies varying the types of the semantic mapping (semantic mapping, semantic feature analysis, and semantic/syntactic feature analysis). All studies consistently showed large effects on comprehension knowledge posttests, which measured understanding of the passage contents and concepts (ES = 1.33, 1.47, and 1.78, respectively). In particular, students who used semantic mapping with additional activities such as semantic feature analysis (semantic mapping predicting the relationship among concepts using the relationship matrix) and semantic/syntactic feature analysis (semantic feature analysis predicting the answers for the close-type of sentences) performed better than the students who used semantic mapping. However, the effects of semantic-mapping, semantic feature analysis, and semantic/syntactic feature analysis on comprehension were not maintained over time (ES = 0.67, 0.44, and 0.46, respectively).

DiCecco and Gleason (2002) also compared the effects of GO versus traditional instruction to improve expository text comprehension. Results indicated that GO revealed medium effect on content knowledge test and very low effect on fact quizzes (ES = 0.48 and 0.07, respectively). DiCecco and Gleason implemented student- and teacher-developed GO through the interaction and discussion. The results showed that the students using GO constructed by both teacher and students outperformed the students who used teacher-developed GO on comprehension.
Gersten et al. (2006) made use of video material, GOs in the form of compare-contrast mapping, and supplemental reading materials. They investigated their effects on the teaching of history. Results indicated a high effect for the content interview test ($ES = 1.10$), where students had to orally answer questions in a one-on-one setting. Authors reported a medium effect for written examinations ($ES = 0.75$) and a medium effect on vocabulary matching procedures ($ES = 0.58$).

### Strategy Instruction

Eight of the 14 studies taught students strategies to help unravel the text. Strategies included summarizing content, identifying the main idea, self-monitoring, underlying structure, collaborative strategic reading, and theme identification (Bakken et al., 1997, Studies 1, 2; Boyle, 2010; Calhoon, 2005; Jitendra et al., 2000; Kim et al., 2006; Mastropieri et al., 2001; Wilder & Williams, 2001). Of the eight studies, Calhoon (2005) used only a standardized test; Kim et al. (2006) used posttests and a standardized test together; and the others employed posttests and follow-up tests without standardized tests. Table 2 provides the $ES$s on those measures for each of the studies.

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention</th>
<th>Reading Components</th>
<th>Fidelity of Instruction</th>
<th>Posttests (Effect Size)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Graphic Organizer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boyle (1996)</td>
<td>Cognitive mapping (Direct instruction)</td>
<td>Comprehension, Vocabulary</td>
<td>Whole group (15)</td>
<td>Standardized Test</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Comprehension post ($ES = 1.33$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Comprehension follow-up ($ES = 0.47$)</td>
</tr>
<tr>
<td>Bos &amp; Anders (1990) Study 1</td>
<td>Semantic mapping</td>
<td>Comprehension, Vocabulary</td>
<td>In groups (6–12)</td>
<td>Researcher-developed measures</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Comprehension post ($ES = 1.47$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Comprehension follow-up ($ES = 0.44$)</td>
</tr>
<tr>
<td>Bos &amp; Anders (1990) Study 2</td>
<td>Semantic feature analysis</td>
<td>Comprehension, Vocabulary</td>
<td>In groups (6–12)</td>
<td>Researcher-developed measures</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Comprehension post ($ES = 1.78$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Comprehension follow-up ($ES = 0.40$)</td>
</tr>
<tr>
<td>Bos &amp; Anders (1990) Study 3</td>
<td>Semantic/syntactic feature analysis</td>
<td>Comprehension, Vocabulary</td>
<td>In groups (6–12)</td>
<td>Researcher-developed measures</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Content knowledge ($ES = 0.48$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Fact quizzes ($ES = 0.07$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Content Interview ($ES = 0.75$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Written exam ($ES = 1.10$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Matching Test ($ES = 0.58$)</td>
</tr>
</tbody>
</table>

Table 1 provides the $ES$s on those measures for each of the studies.

Gersten et al. (2006) made use of video material, GOs in the form of compare-contrast mapping, and supplemental reading materials. They investigated their effects on the teaching of history. Results indicated a high effect for the content interview test ($ES = 1.10$), where students had to orally answer questions in a one-on-one setting. Authors reported a medium effect for written examinations ($ES = 0.75$) and a medium effect on vocabulary matching procedures ($ES = 0.58$).

### Strategy Instruction

Eight of the 14 studies taught students strategies to help unravel the text. Strategies included summarizing content, identifying the main idea, self-monitoring, underlying structure, collaborative strategic reading, and theme identification (Bakken et al., 1997, Studies 1, 2; Boyle, 2010; Calhoon, 2005; Jitendra et al., 2000; Kim et al., 2006; Mastropieri et al., 2001; Wilder & Williams, 2001). Of the eight studies, Calhoon (2005) used only a standardized test; Kim et al. (2006) used posttests and a standardized test together; and the others employed posttests and follow-up tests without standardized tests. Table 2 provides the $ES$s on those measures for each of the studies.
# TABLE 2

Critical Factors Incorporated in Studies: (Strategy Instruction)

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention Component</th>
<th>Reading Components</th>
<th>Group Size</th>
<th>Fidelity of Instruction</th>
<th>Posttests (Effect Size)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main idea and Summarization</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bakken, Mastropieri, &amp; Scruggs (1997) Study 2</td>
<td>Paragraph replacement, main idea</td>
<td>Comprehension</td>
<td>Individual instruction</td>
<td>Script/Researcher</td>
<td>Researcher-developed measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calhoon (2005)</td>
<td>Summarization, main idea (Peer-mediated instruction)</td>
<td>Comprehension, Phonetics</td>
<td>In pairs</td>
<td>Script/Teacher</td>
<td>Researcher-developed measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phonology, Morphology English orthography</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mastropieri et al. (2001)</td>
<td>Summarization, main idea, story retelling (Peer- and reciprocal-instruction)</td>
<td>Comprehension, Fluency</td>
<td>In pairs</td>
<td>Script/Teacher</td>
<td>Researcher-developed measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jirendra, Hoppes, &amp; Xin (2000)</td>
<td>Main idea Self-monitoring</td>
<td>Comprehension</td>
<td>In groups (6 to 8)</td>
<td>Script/Researcher</td>
<td>Researcher-developed measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Underlying Structure</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bakken, Mastropieri, &amp; Scruggs (1997) Study 1</td>
<td>Text-structure based strategy</td>
<td>Comprehension</td>
<td>Individual instruction</td>
<td>Script/Researcher</td>
<td>Researcher-developed measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boyle (2010)</td>
<td>Note-taking strategy</td>
<td>Comprehension</td>
<td>Whole group (20)</td>
<td>Script/Researcher</td>
<td>Researcher-developed measures</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wilder &amp; Williams (2001)</td>
<td>Underlying story theme</td>
<td>Comprehension</td>
<td>In groups (7 to 11)</td>
<td>Script/Teacher</td>
<td>Researcher-developed measures</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Mastropieri et al. also demonstrated that the intervention group outperformed the control group on open-ended comprehension test ($ES = 1.14$). In Calhoon’s study, the intervention group received a phonology program (linguistic skills training-phonology [LST]) as well as a reading comprehension program including summarization and main idea strategies; the control group received regular remedial reading instruction (teacher modeling, feedback, and teacher–student discussion). Mastropieri et al. examined the effects of story retelling, questioning, and summarization strategies in comparison with regular whole-class instruction.

Jitendra et al. (2000) examined the effectiveness of a main idea/self-monitoring strategy versus regular reading instruction in the resource room. The intervention revealed high effects on the posttests including training, near-transfer, and far-transfer subtests ($ES = 3.40, 1.95$ and $1.45$, respectively) and a total posttest ($ES = 2.20$). However, the effect of strategy decreased on follow-up tests including training, near-transfer, and far-transfer sub-tests ($ES = 2.00, 0.57$ and $0.77$, respectively) with a total follow-up test ($ES = 0.99$), compared to the results in the posttests.

One study, Kim et al. (2006), investigated the effects of computer-assisted collaborative strategic reading (CACSR) on reading comprehension. The intervention group that received CACSR demonstrated a high effect on the post main idea test ($ES = 0.76$); however, the effect was less on the standardized test, Woodcock Reading Mastery Test–Revised Passage Comprehension subtest ($ES = 0.49$). Collaborative strategic reading (CSR) consists of a set of strategies based on the reciprocal teaching model that requires students to predict, question, monitor, and review their reading. The control group received regular reading instruction including vocabulary, fluency, and comprehension.

The remaining three studies (Bakken et al., 1997, Study I; Boyle, 2010; Wilder & Williams, 2001) that employed strategy instruction targeted the underlying structure and concepts of the text and evaluated the results using posttests and follow-up tests. Results of all three studies were in favor of treatment groups (range $ES = 0.49–5.23$). Bakken et al. investigated the effects of underlying text-structure strategy versus traditional instruction on comprehension (content recall) using science and social texts. The strategy was successful in enhancing comprehension abilities of students with LD (range $ES = 2.14–3.12$). In the intervention group, students were taught how to identify structures of different types of passages (main idea, list, and order passages) and organize the passages (see Table 3 for example passages).

Boyle (2010) investigated the effects of strategic note-taking strategy on content recall. The intervention group outperformed the control group, demonstrating high effects on posttests (range $ES = 0.83–1.01$) and follow-up tests (range $ES = 0.83–0.88$). Students in the intervention group were provided a strategic note sheet to write main points.

### Table 2

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention Components</th>
<th>Group Size</th>
<th>Posttests (Effect Size)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kim et al. (2006)</td>
<td>Predicting, summarizing, monitoring, and questioning (Computer assisted instruction)</td>
<td>In pairs</td>
<td>Standardized test</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nonscript/Teacher</td>
<td>- WRMT-R passage comprehension (0.48)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Instructional level Gist (main idea) (0.76)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Instructional level multiple-choice (0.83)</td>
</tr>
</tbody>
</table>

### Table 3

**Examples of Main Idea, List, and Order Passages**

**Main Idea**
- A passage consists of main idea and supporting sentences

**List**
- A passage consists of main topic and list of subtopics

**Order**
- A passage consists of main topic and subtopics following specific steps

**Example of Passages**

**Main Idea**
- Acid rain is an increasing problem as more factories produce goods in the United States. Chemicals are released from smokestacks of the factories into the air. Those chemicals combine with water vapor to form acids that fall on the earth as rain, snow, or fog. Studies show that by decreasing the amount of chemicals in the air the acid rain can be reduced.

**List**
- Water that doesn’t soak into the ground or evaporate flows across the Earth’s surface and is called runoff. Runoff can be affected by things such as: the amount of rain; the time span during which the rain falls; the slope of the land; and the amount of vegetation on the land, such as grass.

**Order**
- Sedimentary rocks come from other rocks. First, rocks are broken into smaller pieces. Second, pieces are moved by water, wind, ice, or gravity. Third, erosion moves these pieces to a new location where they are left, and layer upon layer builds up. Finally, pressure from the upper layers makes down on the lower layers to form sedimentary rocks.
important vocabularies and a summary of the lecture, while the control group was provided a blank sheet.

Wilder and Williams (2001) examined the effects of underlying theme story strategy in improving comprehension abilities. The strategy indicated high effects in theme identification tests on a narrative text with and without prompt (range $ES = 1.19-5.23$), but medium effect for generating a story incorporating the theme ($ES = 0.49$). Students in the intervention group learned how to identify a story theme using a step-procedure (identification of a theme, generalization of the theme, and then application of the generalized theme), and the comparison group received regular instruction (e.g., reading the story, questions, and discussion).

Summary

The findings of this review indicate that instructional modifications are helpful in improving comprehension skills, but their effects vary. For example, some studies, using GOs resulted in high $ES$s, while other studies demonstrated little or no difference between intervention and control groups. Additionally, the results on follow-up tests demonstrated that the effects were not maintained. On the other hand, students who used strategies consistently, revealed high comprehension scores. In particular, main idea, summarization, and targeting underlying structures were more beneficial than the other strategies for students. The $ES$s were high on posttests as well as on the standardized test.

Effects by Self-Monitoring

We examined effects of self-monitoring on comprehension performance. Of the 14 studies, only Jitendra et al. (2000) conducted self-monitoring. Jitendra et al. examined the relative effects of multiple-strategy instruction (self-monitoring in addition to main idea) versus traditional reading instruction. Students in the intervention group scored significantly higher on total posttest and total follow-up test ($ES = 2.20$ and 0.99, respectively). Specifically, the performance of the intervention group was higher on main idea posttest scores ($ES = 3.40$) than near-transfer ($ES = 1.95$) and far-transfer ($ES = 1.43$) tests. Treatment effects were still high for the main-idea follow-up test administered 6 weeks after instruction concluded ($ES = 2.00$), but not as high for near-transfer ($ES = 0.57$) or far-transfer tests ($ES = 0.77$).

To examine the effect of self-monitoring alone from among multiple strategies targeting comprehension, we compared results of the study that used main idea with self-monitoring to the results of the studies that used main idea without self-monitoring. Three studies (Bakken et al., 1997, Study 2; Calhoon, 2005; Mastropieri et al., 2001) used main idea identification without self-monitoring, and students who received the strategy consistently performed high on posttests (range $ES = 0.84-1.41$, mean $ES = 1.15$) and on the standardized test ($ES = 0.84$). However, the effects of main idea in combination with self-monitoring (Jitendra et al., 2000) were higher on posttests (range $ES = 1.43-3.40$, mean $ES = 2.26$), indicating a possible additional effect of self-monitoring in improving comprehension.

Effects by Reading Components Incorporated

We reviewed studies for the reading components they incorporated and their possible impact on reading outcomes. We organized studies based on whether they taught comprehension alone (focused instruction) or multiple reading components together (balanced instruction). Eight of 14 studies implemented focused instruction (Boyle, 1996, 2010; Bakken et al., 1997; Studies 1, 2; Jitendra et al., 2000; Gersten et al., 2006; Kim et al., 2006; Wilder & Williams, 2001). The results of analysis indicated that the effects of focused instruction on comprehension varied. The $ES$s were medium to high on posttests (range $ES = 0.49-5.23$) and lower on standardized tests (range $ES = 0.34-0.48$). In the other six studies, comprehension was taught in combination with other components of reading (Bos & Anders, 1990, Studies 1, 2, 3; Calhoon, 2005; DiCecce & Gleason, 2002; Mastropieri et al., 2001). The effects of balanced instruction were inconsistent on posttests (range $ES = 0.07-1.78$) and on the standardized test ($ES = 0.84$) just as it was in focused instruction.

Assuming that the type of instructional methods influenced the results, we compared the $ES$s in the studies that used focused instruction with those that used balanced instruction after grouping the studies into instructional modifications or strategy instruction. When instructional modifications were used, the effects of the focused instruction (Boyle, 1996; Gersten et al., 2006) ranged between small and large on posttests (range $ES = 0.58-1.33$) and was small on the standardized test ($ES = 0.34$). Also, for balanced instruction (Bos & Anders, 1990, Studies 1, 2, 3; DiCecce & Gleason, 2002), the $ES$s ranged from small to large on posttests (range $ES = 0.07-1.78$) and small on follow-up tests (range $ES = 0.40-0.47$). Among the balanced instruction studies that targeted instructional modifications, instruction that incorporated vocabulary (Bos & Anders, 1990, Studies 1, 2, 3) yielded consistently higher effects (mean $ES = 1.53$) than instruction that incorporated vocabulary and word-decoding (DiCecce & Gleason, 2002) on posttests (mean $ES = 0.28$).

Of the studies that used strategy instruction for main idea and summarization (Bakken et al., 1997, Study 1; Calhoon, 2005; Jitendra et al., 2000; Mastropieri et al., 2001), the effects of intervention were large regardless of the reading components they incorporated. The range of $ES$s was from 1.14 to 1.75 on posttests; the $ES$ on the standardized test was 0.84. All studies that used underlying structure (Bakken et al., 1997, Study 1; Boyle, 2010; Wilder & Williams, 2001) employed focused instruction, and students performed high on all posttests and follow-up tests (range $ES = 0.83-3.12$). Kim et al. (2006) also employed focused instruction CACSR; the effects were medium to large on the posttests (range $ES = 0.76-0.83$) and medium on the standardized test ($ES = 0.48$).

We hypothesized that students would have needed differentiated instruction based on their prior knowledge. For example, if students perform low on vocabulary and
comprehension, they would need balanced instruction incorporating vocabulary and comprehension. Thus, we identified students' pretest outcomes on reading in each study and reanalyzed data, but the results were inconclusive. The students in the study of Gersten et al. (2006) performed low on rapid automatized naming (Denckla & Rudel, 1976) and oral reading fluency (Shinn, Good, Knutson, Tilly, & Collins, 1992). They employed focused instruction and ESs were medium to high on posttests (range ES = 0.58–1.10). The students in the study of Boyle (1996) also performed low on multiple reading components including comprehension and word attack (Stanford Diagnostic Reading Test). However, in this study, focused instruction yielded high effects in posttests (range ES = 0.96–1.23). The studies that used balanced instruction (Bos & Anders, 1990, Studies 1, 2, 3) did not specify students' vocabulary abilities on pretests, so we were not able to analyze the data.

In summary, it appeared that the effects of instruction on comprehension varied regardless of the reading components they incorporated. Although this body of literature is not comprehensive, when providing balanced instruction, results do point toward the beneficial effects of specifically incorporating vocabulary versus other components in reading comprehension instruction.

**Effects by Fidelity of Instruction**

We investigated the effects of fidelity of instruction delivery. We measured fidelity by examining whether an instructor used a script or not. Ten of the 14 studies used a script (Bakken et al., 1997, Studies 1, 2; Boyle, 1996, 2010; Calhoon, 2005; DiCecco & Gleason, 2002; Gersten et al., 2006; Jitendra et al., 2000; Mastropieri et al., 2001; Wilder & Williams, 2001). The effects of the use of a script varied across studies, with ESs ranging from small to large on posttests (range ES = 0.07–5.23) and on standardized tests (range ES = 0.34–0.84). The other four studies did not employ scripted instruction (Bos & Anders, 1990, Studies 1, 2, 3; Kim et al., 2006). The range of ESs in nonscripted instruction was medium to high on posttests (range ES = 0.76–1.78) and low on the standardized test (ES = 0.48).

We analyzed the results of fidelity of instruction by adding another variable: researcher-delivered instruction (instruction delivered by authors) or teacher-delivered instruction (instruction delivered by classroom teacher of the participating students). Thus, interventions were delivered in one of four ways: researcher-delivered instruction with script, researcher-delivered instruction with no script, teacher-delivered instruction with script, and teacher-delivered instruction with no script. In five studies, interventions were delivered by researchers with script (Bakken et al., 1997, Studies 1, 2; Boyle, 1996, 2010; Jitendra et al., 2000). The studies indicated high effects across posttests (range ES = 0.86–3.40), but medium to high effects on follow-up tests (range ES = 0.57–3.12). A second group of studies (Bos & Anders, 1990, Studies 1, 2, 3) delivered interventions by researcher with no script. This second group of studies reported high effects on posttests (range ES = 1.33–1.78), but smaller effects on follow-up tests (range ES = 0.40–0.47).

In the other six studies, the interventions were conducted by trained teachers; they reported mixed effects across studies and the type of measure. Five of the six studies used scripted instruction (Calhoon, 2005; DiCecco & Gleason, 2002; Gersten et al., 2006; Mastropieri et al., 2001; Wilder & Williams, 2001). Calhoon (2005) and Mastropieri et al. (2001) reported high effects on the standardized test and posttest (ES = 0.84 and 1.14, respectively). Still, Gersten et al. (2006) reported mixed effects across measures (range ES = 0.58–1.10) and DiCecco and Gleason (2002) indicated small effects on the posttests, content knowledge test and fact quiz (ES = 0.48 and 0.07, respectively). In Kim et al. (2006), the students were taught by trained teachers without scripts and the results varied across measures.

Overall, the findings showed that the effects of interventions were consistently high when delivered by a researcher, relative to instruction delivered by a teacher. The effect sizes for studies incorporating teacher-delivered instruction, ranged from small to high.

**Effects by Group Size**

We examined the effects by group size on comprehension performance. The studies were grouped into four based on the group size: whole group instruction (12 or more students), instruction in groups (6–12 students), instruction in pairs, and individual instruction. None of the studies conducted small group instruction (three to five students); we could not examine the effects of small group instruction.

Of the 14 studies reviewed, three (Boyle, 1996, 2010; DiCecco & Gleason, 2002) implemented whole group instruction; students in five studies (Bos & Anders, 1990, Studies 1, 2, 3; Jitendra et al., 2000; Wilder & Williams, 2001) were instructed in groups; students in four studies (Calhoon, 2005; Gersten et al., 2006; Kim et al., 2006; Mastropieri et al., 2001) learned in pairs; and students in two studies (Bakken et al., 1997, Studies 1, 2) were individually instructed. The analysis revealed that the individual instruction was most effective on posttests (range ES = 1.20–2.57) and follow-up tests (range ES = 1.23–3.12). Except for the studies that employed individual instruction, the effects were inconsistent for the other three groups (whole group, groups, and pairs) across studies and measures. The range of ESs was small to large on posttests (range ES = 0.07–5.23), follow-up tests (range ES = 0.40–2.00), and standardized tests (range ES = 0.34–0.84).

For a more systematic analysis, we analyzed the effects of group size after categorizing the studies by instruction types (instructional modifications and strategy instruction). In the studies that use instructional modifications (Boyle, 1996; DiCecco & Gleason, 2002), ESs ranged from low to high on posttests (range ES = 0.07–1.33) for students instructed as a whole group. Students instructed in groups (Bos & Anders, 1990, Studies 1, 2, 3) performed high on posttests (range
As regards selecting a grouping format, this synthesis suggests that teachers look for alternatives to whole group instruction. Results of this synthesis indicated that smaller group instruction (6–12 students) yielded higher $ES$s than whole group (12 or more) or paired instruction, particularly when using instructional modifications. For interventions employing strategy instruction, the review noted that paired instruction seemed to be implemented more effectively with main idea and summarization strategy. Calhoon (2005) and Mastropieri et al. (2001) used summarization and main idea strategy with PALS or a peer-tutoring model. Both studies revealed high $ES$s on standardized measure and comprehensive comprehension measure. This result indicates that when we use main idea and summarization, a paired model of instruction may be beneficial. Strategy instruction also yielded high effects when used with
individualized instruction. Therefore, when employing instructional modifications, teachers could consider instructing in smaller groups, of 6–12 students, while for strategy instruction, instructing in pairs or individually may be more beneficial. It is important to note that none of the studies in this synthesis employed small group instruction (three to five students). This grouping format has been seen to be the most effective with elementary students (Elbaum et al., 2000). However, we could not examine the effect of the small group instruction grouping format.

Limitations

Results of this review should be viewed in the light of several limitations. First, this synthesis investigated the effects on comprehension by group size, but not duration. We initially did seek to analyze studies based on the intervention duration. However, durations of the studies examined were too varied to draw any conclusive findings. Duration of instruction in the studies reviewed ranged from a minimum of 3 days to maximum 12 weeks. Also, the frequency per week and duration per session varied significantly across studies. A second limitation is that we could not show the effect of small group (groups of three to five) instruction when examining results by type of group. None of the studies reviewed implemented small group instruction, a common grouping format in special education (Vaughn et al., 2003). Finally, we included studies that included only middle school students as their participants. There are more than 50 studies investigating the effects of intervention on comprehension, published from 1990 to 2010. However, the studies satisfied that the criteria were limited. Many studies included either middle school and high school students, or middle school students with upper elementary school students. Future research may indicate more conclusive results by broadening the search criteria.

Implications for Practice

The purpose of this synthesis was to guide teachers when making instructional decisions regarding reading comprehension instruction. Despite the limitations discussed above, this review has several implications for practitioners.

First, with regard to choosing among instructional strategies, this review suggests that teachers make preferential use of strategy instruction rather than instructional modifications. Although both strategy instruction and instructional modifications are associated with positive results, strategies such as main idea generation, summarization and identifying the underlying structure of the text were seen to more impact reading comprehension as compared to instructional modifications such as GOs and supplemental materials.

When adopting strategy instruction, this synthesis highlights the importance of incorporating a self-monitoring component. In particular, self-monitoring and the main idea strategy were associated with high effects.

Next, when identifying what aspects of reading to incorporate, this synthesis suggests that focused instruction, and balanced instruction are both beneficial. If teachers elect to use balanced instruction, vocabulary in particular of the five reading components, was associated with higher effect sizes, and should therefore be incorporated.

When selecting a grouping format, this synthesis suggests that teachers look for alternatives to whole group instruction which was seen to be least effective among the various grouping formats. When employing instructional modifications, teachers could consider instructing in groups (6–12 students), and for strategy instruction, instructing in pairs of individually.

Finally, this synthesis suggests that teachers keep abreast of why they are employing a particular strategy and spend time understanding how best to deliver the intervention. Teacher educators need to emphasize the content of the intervention as well as delivery when providing professional development training.

REFERENCES


About the Authors

Woori Kim is Postdoctoral Fellow at Seoul National University. Her research interests include Tiered instruction and reading interventions for students with learning disabilities.

Sylvia Linan-Thompson is an associate professor at The University of Texas at Austin. Her research interests include examining the effectiveness of interventions and assessments for English language learners and the implementation of RTI.

Radhika Misquitta is the Director – Upper School of the Gateway School of Mumbai. Her research interests include reading and mathematics instruction for older students with disabilities.