Graphic Organizers and Their Effects on the Reading Comprehension of Students with LD: A Synthesis of Research

Ae-Hwa Kim, Sharon Vaughn, Jeanne Wanzek, and Shangjin Wei

Abstract

Previous research studies examining the effects of graphic organizers on reading comprehension for students with learning disabilities (LD) are reviewed. An extensive search of the professional literature between 1963 and June 2001 yielded a total of 21 group design intervention studies that met the criteria for inclusion in the synthesis. Using graphic organizers (i.e., semantic organizers, framed outlines, cognitive maps with and without a mnemonic) was associated with improved reading comprehension overall for students with LD. Compared to standardized reading measures, researcher-developed comprehension measures were associated with higher effect sizes. Initial gains demonstrated when using graphic organizers were not revealed during later comprehension tasks or on new comprehension tasks.

Graphic organizers are visual and spatial displays designed to facilitate the teaching and learning of textual material through the “use of lines, arrows, and a spatial arrangement that describe text content, structure, and key conceptual relationships” (Darch & Eaves, 1986, p. 310). Graphic organizers include semantic maps, semantic feature analysis, cognitive maps, story maps, framed outlines, and Venn diagrams.

Ausubel (1963) originally rationalized the use of graphic organizers by speculating that a learner’s existing knowledge, which he referred to as cognitive structure, greatly influences his or her learning. When the cognitive structure expands and strengthens by incorporating new information, learning occurs. To facilitate this process, graphic organizers provide learners with a meaningful framework for relating their existing knowledge to the new information (Ausubel, 1963; Wittyrock, 1992). Mayer (1984) reaffirmed the utility of graphic organizers by examining reading as an information processing and storage process during which graphic organizers may be used to display connections among concepts. Within the reading process, Mayer suggested that the use of such organizers may better allow readers to connect their existing knowledge base with the text information.

By using graphic organizers, educators hope to facilitate the readers’ understanding of the text through visual depictions of key terms and concepts and the relationships among them (Simmons, Griffin, & Kame’enui, 1988). Of course, enhancing all students’ reading comprehension is valuable, but making such improvements is particularly important and demanding when focusing on students with learning disabilities (LD). As students progress through school, reading increasingly involves expository text from which students are expected to learn; however, this type of text may pose more challenges than other reading assignments (e.g., fiction). For example, expository text is more information driven and may contain unfamiliar technical vocabulary. Moreover, such text is often organized poorly and, consequently, requires students to perform complex cognitive tasks in order to comprehend the material (Lapp, Flood, & Ranck-Buhr, 1995). Many students with LD who may already have difficulty with reading and study skills are challenged by the academic domains of reading and interpreting expository text (Bryant, Ugel, Thompson, & Hamff, 1999; Bryant et al., 2000). Therefore, these students can benefit from learning strategies that facilitate the comprehension of this type of text. One strategy that has often been recommended to assist students with LD in learning from expository text is the use of graphic organizers (Bos & Vaughn, 2002; Rivera & Smith, 1997; Taylor, Harris, & Pearson, 1988).
Although practitioner journals or texts recommend the use of graphic organizers as an effective strategy to promote the comprehension of expository text (Blachowicz & Ogle, 2001; Boyle & Yeager, 1997; Hudson, Lignugaris-Kraft, & Miller, 1993), the accumulated body of research demonstrates inconclusive findings for supporting such a recommendation. Over the past 3 decades, researchers have conducted several syntheses of the research literature to analyze how the use of graphic organizers with expository text affects students’ comprehension (Griffin & Tulbert, 1995; Moore & Readence, 1980, 1984). In general, these syntheses revealed that using graphic organizers has small overall effects on students’ comprehension of expository text. For example, Moore and Readence (1980, 1984) conducted two meta-analyses to examine the effects of graphic organizers on the comprehension of expository text. Their findings revealed small overall effect sizes (.30 and .22, respectively), which suggests a somewhat limited value of using graphic organizers to promote comprehension.

In addition to these earlier analyses, Griffin and Tulbert (1995) reviewed several studies relating the use of graphic organizers to reading comprehension that yielded inconclusive findings, which they attributed to variations in intervention designs and instructional processes and to confounds with methodology. Similarly, Rice (1994) argued that the lack of consistent operational criteria in studies using graphic organizers contributed to inconclusive findings. For example, instructional procedures for using graphic organizers varied by such factors as text variables (e.g., type, content, level), originators (e.g., teachers vs. students), position relative to the text (e.g., before or after the text), and outcome measures. Although several reviews provided tenuous results on the differential effects of various graphic organizers—using graphic organizers constructed by students, positioned after the text, and used for a longer time demonstrated greater effects than using graphic organizers constructed by teachers, positioned before the text, and used for a shorter time (Dunston, 1992; Moore & Readence, 1980, 1984)—overall, the effects of using graphic organizers still appear to be inconclusive due to the variations in operational criteria in studies using graphic organizers.

In general, previous syntheses focused on examining the effect of using graphic organizers on the reading comprehension of students without LD. The one review that examined the effect of using graphic organizers specifically on the reading comprehension of students with LD (Griffin & Tulbert, 1995) only included three studies. Because of this limited number of studies, it is not possible to reach definitive conclusions from the findings for students with LD.

Although students with LD have difficulty organizing and recalling verbal information (Wong, 1978), they may perform nonverbal tasks fairly successfully (Vellutino, Harding, Stager, & Phillips, 1975). Similarly, researchers have argued that students with LD have strengths in spatial or visual modes of conceptualization (Pirozzolo & Rayner, 1979; Witelson, 1977). Thus, one can hypothesize that visual displays of information, such as graphic organizers, may help students with LD circumvent their difficulties with organizing and recalling verbal information, thereby enhancing their reading comprehension. To determine how effective the use of graphic organizers is for enhancing the reading comprehension of students with LD, a systematic synthesis of research is needed. Although most research on graphic organizers has focused on students without LD, quite a few research studies have been conducted specifically with students with LD.

The purpose of this article is to systematically review the findings of group design intervention studies examining the effect of graphic organizers on comprehension for students with LD. This synthesis is intended to assist practitioners and future researchers to better understand the effects of using graphic organizers to improve the reading comprehension of students with LD.

Method

Study Selection

We used a three-step process to identify studies to include in the synthesis. First, journal articles and dissertations that met specific criteria were selected from a meta-analysis by Swanson, Hoskyn, and Lee (1999), in which they systematically scanned online databases (i.e., PsycINFO, MEDline, ERIC, and Dissertation Abstracts) for studies from 1963 to 1997. Criteria used to select the studies for our synthesis included the following:

1. Students in K–12. Participating students must have been in Grades K–12.
2. Identified disability. Participants must have included students with LD.
3. Design. Research design must have been either a treatment–comparison design or a single-group design. Single-subject design studies were not included; however, we searched for such studies during a hand search of 10 journals and present these studies’ findings in the discussion.
4. Independent variable. Intervention must have been the use of graphic organizers (i.e., semantic feature analysis, semantic maps, or other visual organizers) either displaying concept relationships discussed within the text or providing an outline/overview of the text. Graphic organizers not serving one of these purposes were not included (e.g., Bergerud, Lovitt, & Horton, 1988).
5. Dependent variable. Reading comprehension must have been one of the dependent measures, and this measure had to assess students’ silent or oral
reading and their ability to answer questions about the passage. For example, several studies conducted by Bos and colleagues included a concept measure, which assessed students’ ability to apply the concepts presented in the passage to real-life situations (Anders, Bos, & Filip, 1984; Bos & Anders, 1990, 1992; Bos, Anders, Filip, & Jaffe, 1985, 1989). The concept measure included passages for the students to read and demonstrate their understanding by answering questions; therefore, we considered this instrument a reading comprehension measure.

After the initial search, we conducted a second set of computer searches in PsycINFO and ERIC to locate studies published between 1996 and June 2001 that also met these criteria. Key words or root forms of them (i.e., reading comprehension, comprehension, reading, learning disab*, reading disab*, dyslex*) were entered in various combinations to capture the greatest possible number of articles. A final search was conducted by hand in 10 major journals related to the topic from 1996 through June 2001. Journals used in this final search included Applied Psycholinguistics, Exceptional Children, Journal of Educational Psychology, Journal of Learning Disabilities, The Journal of Special Education, Learning Disability Quarterly, Learning Disabilities Research & Practice, Remedial and Special Education, Reading Research Quarterly, and Reading and Writing.

**Data Analysis**

**Coding Procedures.** Through the use of extensive coding, we organized the pertinent information from each intervention study, including age of participants, stated purpose, research design, intervention variables, and reported findings. After this information was double-coded for all of the articles, an interrater reliability of .96 (range = .75 to 1.00) was found. Furthermore, the first author reviewed all code sheets for comprehensiveness and accuracy, and for the few cases containing questionable codes, the researchers met to resolve the ambiguities and reach a decision by consensus.

**Effect Size Calculation.** For intervention studies with sufficient statistical information, we calculated the effect size, $d$. For treatment–comparison design studies, the effect size was calculated as the difference between the intervention group’s mean posttest score and the comparison group's mean posttest score divided by the pooled standard deviation. A standardized mean change measure, on the other hand, was used to calculate effect sizes for the single-group design studies (Becker, 1988). When the researchers reported only a $t$ for an $F$ statistic, we estimated effect size by applying the following formulas derived by Rosenthal and colleagues (Rosenthal, 1991; Rosenthal & Rosnow, 1984):

$$d = \frac{2t}{\sqrt{df}}$$

$$d = \frac{2\sqrt{F}}{\sqrt{df(error)}}$$

When authors reported statistically nonsignificant results with no additional information, we assigned 0 as the effect size (Bos et al., 1985, 1989). Assigning 0 to nonsignificant results has a conservative impact on the results (Cooper, 1998). We calculated effect sizes for all but one study (i.e., Boyle & Weishaar, 1997). We interpreted effect sizes according to Cohen’s (1988) guidelines: 0.2 as a small effect size, 0.5 as a medium one, and 0.8 as a large one.

**Results**

A total of 21 intervention studies, which were all reported in journal articles, met our criteria for inclusion in the synthesis. Six of these studies were reported in one article (Bos & Anders, 1992), whereas another journal article contained three studies (Horton, Lovitt, & Bergerud, 1990). Moreover, the same two studies were reported in more than one journal (Bos et al., 1985, 1989). One article was excluded from our synthesis because, although the use of graphic organizers was mentioned in the abstract, the authors did not include such an implementation in either the intervention’s description or the results or discussion sections (Weisberg & Balajthy, 1990). These results led us to review a total of 15 journal articles containing 21 separate intervention studies. Each study’s purpose, participants, interventions, measures, and findings (i.e., effect sizes or descriptive findings) are described in Table 1 for the treatment–comparison design studies and in Table 2 for the single-group design studies.

**Features of the Intervention Studies**

All of the studies included in this synthesis reported on interventions using graphic organizers to improve the reading comprehension of students with LD. In the following sections, we describe the interventions’ features including participants, research designs, intervention variables, and comprehension measures.

**Participants**

The 21 studies included a total of 848 students with LD. Moreover, 16 students with educable mental retardation were included in two of the studies (Boyle, 1996, 2000). Four studies also included students without disabilities ($n = 724$; Horton et al., 1990, Studies 1, 2, & 3; Lovitt, Rudsit, Jenkins, Pious, & Benedetti, 1986); however, when a study included such students, the findings were disaggregated for students with LD (Horton et al., 1990, Studies 1, 2, & 3; Lovitt et al., 1986). Of the 21 studies, 6 included high school students, another 6 included junior high school students, and 5 included elementary school students. Furthermore, both elementary and junior high school students participated in one
<table>
<thead>
<tr>
<th>Study/Participants</th>
<th>Intervention</th>
<th>Dependent measures</th>
<th>Key findings (effect size)</th>
</tr>
</thead>
</table>
| Anders, Bos, & Filip (1984) 62 high school students with LD | • T (Semantic feature analysis): A relationship matrix with the major concepts and vocabulary from the passages.  
• C (Dictionary instruction): Looking up difficult words in the dictionary, writing their definitions, and writing a sentence using each word.  
• Duration/Intensity: Two 50-minute practice sessions and two 50-minute experimental sessions over 2 weeks. | • Researcher-developed multiple-choice comprehension test | T vs. C (posttest): \( ES = 1.52, p < .001 \) |
| Bos & Anders (1990) 61 junior high students with LD | • T1 (Semantic mapping): A hierarchical relationship map from the vocabulary list.  
• T2 (Semantic feature analysis): A relationship matrix with the major concepts and vocabulary from the passages.  
• T3 (Semantic/syntactic feature analysis): Same as T2, but in this condition, using the matrix to answer cloze-type sentences.  
• C (Definition instruction): Direct instruction on the definitions of the vocabulary terms.  
• Duration/Intensity: Three practice sessions and three 50-minute experimental sessions over 2 weeks. | • Researcher-developed multiple-choice comprehension test | T1 vs. C (posttest): \( ES = 1.31, p < .05 \)  
T2 vs. C (posttest): \( ES = 1.49, p < .05 \)  
T3 vs. C (posttest): \( ES = 1.22, p < .05 \)  
T1, T2, T3 vs. C (posttest): \( ES = 0.81^a \) |
| Bos & Anders (1992) Study 1 42 bilingual elementary students with LD | Phase I: Researchers implemented interventions in whole groups.  
• T1 (Semantic mapping)  
• T2 (Semantic feature analysis)  
• T3 (Semantic/syntactic feature analysis)  
• C (Definition instruction)  
• Duration/Intensity: Three 50-minute practice sessions and three 50-minute experimental sessions over 2 weeks. | • Researcher-developed multiple-choice comprehension test | T1, T2, T3 vs. C (posttest): \( ES = 0.81^a \)  
T1, T2, T3 vs. C (follow-up): \( ES = 0.86^a \) |
| Bos & Anders (1992) Study 2 61 junior high students with LD | Phase I: Researchers implemented interventions in whole groups.  
• Same as Study 1 | • Researcher-developed multiple-choice comprehension test | T1, T2, T3 vs. C (posttest): \( ES = 1.22^a \)  
T1, T2, T3 vs. C (follow-up): \( ES = 0.78^a \) |
| Bos & Anders (1992) Study 3 47 bilingual elementary students with LD | Phase II: Special education teachers implemented interventions in whole groups.  
• T1 (Semantic mapping)  
• T2 (Semantic feature analysis)  
• T3 (Semantic/syntactic feature analysis)  
• C (Definition instruction)  
• Duration/Intensity: Three 50-minute practice sessions and three 50-minute experimental sessions over 2 weeks | • Researcher-developed multiple-choice comprehension test | T1, T2, T3 vs. C (posttest): \( ES = 1.46^a \) |
| Bos & Anders (1992) Study 4 53 junior high students with LD | Phase II: Special education teachers implemented interventions in whole groups.  
• Same as Study 3 | • Researcher-developed multiple-choice comprehension test | T1, T2, T3 vs. C (posttest): \( ES = 1.51^a \)  
T1, T2, T3 vs. C (follow-up): \( ES = 1.51^a \) |
| Bos & Anders (1992) Study 5 26 bilingual elementary students with LD | Phase III: Special education teachers implemented interventions in small cooperative learning groups. They acted as facilitators and coaches.  
• T1 (Semantic mapping) | • Researcher-developed multiple-choice comprehension test | T1 vs. C (posttest): \( ES = 0.48^a \)  
T1 vs. C (follow-up): \( ES = 0.64^a \) |

(table continues)
### Table 1 continued

<table>
<thead>
<tr>
<th>Study/Participants</th>
<th>Intervention</th>
<th>Dependent measures</th>
<th>Key findings (effect size)</th>
</tr>
</thead>
</table>
| Bos & Anders (cont’d) | • T2 (Semantic feature analysis)  
  • C (Normative group): Average-achieving students who read same passages.  
  • Duration/Intensity: Three 50-minute practice sessions and three 50-minute experimental sessions over 2 weeks. | T2 vs. C (posttest): $ES = 0.31^a$  
  T2 vs. C (follow-up): $ES = 0.06^a$ | |
| Bos & Anders (1992) Study 6 | Phase III: Special education teachers implemented interventions in small cooperative learning groups. They acted as facilitators and coaches.  
  • Same as Study 5 | • Researcher-developed multiple-choice comprehension test | T1 vs. C (posttest): $ES = 2.25^a$  
  T1 vs. C (follow-up): $ES = 0.81^a$  
  T2 vs. C (posttest): $ES = 0.18^a$  
  T2 vs. C (follow-up): $ES = 0.76^a$ |
| Bos, Anders, Filip, & Jaffe (1985); Bos, Anders, Filip, & Jaffe (1989) 50 high school students with LD | • T (Semantic feature analysis)  
  • C (Dictionary instruction)  
  • Duration/Intensity: Four 50-minute experimental sessions over 2 weeks. | • Researcher-developed multiple-choice comprehension test | T vs. C (posttest): $ES = 1.69$, $p < .001$  
  T vs. C (follow-up): $ES = 0$, ns |
| Boyle (1996) 30 sixth, seventh, and eighth graders with LD or EMR (20 with LD) | • T (Cognitive mapping strategy–TRAVEL):  
  Topic–Write down the topic and circle it.  
  Read–Read a paragraph.  
  Ask–Ask what the main idea and three details are and write them down.  
  Verify–Verify the main idea by circling it and linking its details.  
  Examine–Examine the next paragraph and ask and verify again.  
  Link–When finished with the story, link all of the circles.  
  • C: Typical reading techniques.  
  • Duration/Intensity: Six 50-minute sessions over two weeks. | • Formal Reading Inventory: Comprehension  
  • CB: Below grade level, literal  
  • CB: Below grade level, inferential  
  • CB: On grade level, literal  
  • CB: On grade level, inferential  
  • SDRT | T vs. C (posttest): $ES = 0.35$, ns  
  $ES = 0.89$, $p < .05$  
  $ES = 0.78$, $p < .05$  
  $ES = 1.37$, $p < .05$  
  $ES = 0.98$, $p < .05$  
  $ES = 0.48$, ns |
| Boyle (2000) 24 ninth and tenth graders with LD or EMH (18 with LD) | • T (Cognitive mapping strategy–RELATE):  
  Read quickly to locate each topic.  
  Etch out one circle for each topic.  
  Look for ideas unique to each topic and those related to more than one topic as you start to read.  
  Anchor three unique ideas for each topic in each circle.  
  Tie together two or three ideas related to two of the topics.  
  Enclose three ideas related to all three topics.  
  • C: Typical reading techniques.  
  • Duration/Intensity: Two 50-minute sessions over 1 week. | • CB with two-topic passage, literal  
  • CB with two-topic passage, inferential  
  • CB with three-topic passage, literal  
  • CB with three-topic passage, inferential  
  • CB with three-topic passage, rational | T vs. C (posttest): $ES = 1.22$, $p < .05$  
  $ES = 0.78$, ns  
  $ES = 1.03$, $p < .05$  
  $ES = 1.18$, $p < .05$  
  $ES = 0.36$, ns  
  $ES = 0.87$, $p < .05$ |
| Boyle & Weishaar (1997) 39 tenth, eleventh, and twelfth graders with LD | • T1 (Student-generated cognitive organizer–TRAVEL): Students learned the cognitive mapping strategy TRAVEL (see Boyle, 1996).  
  • T2 (Expert-generated cognitive organizer–TRAVEL): Students used organizers created by researcher-developed multiple-choice comprehension test  
  • Formative Reading Inventory: Comprehension  
  • CB: Below grade level, literal, Below grade level, inferential, On grade level, literal, On grade level, inferential  
  • SDRT | T vs. T2 vs. C (posttest): $T1 > C$ on all four measures, $p < .05$  
  $T2 > C$ on one of four measures (below grade level literal), $p < .05$ | |

*Table continues*
### Table 1 continued

<table>
<thead>
<tr>
<th>Study/Participants</th>
<th>Intervention</th>
<th>Dependent measures</th>
<th>Key findings (effect size)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Boyle &amp; Weishaar (cont’d)</strong></td>
<td>the researchers using the TRAVEL strategy to guide them during reading.</td>
<td>• T1 &gt; T2 group on one of four measures (below grade level, literal), p &lt; .05</td>
<td>Overall ns</td>
</tr>
<tr>
<td></td>
<td>• C: Typical reading techniques.</td>
<td>T vs. C (posttest): ES = 1.79, p &lt; .01</td>
<td>T vs. C (transfer test): ES = 0.67, ns</td>
</tr>
<tr>
<td></td>
<td>• Duration/Intensity: Eight 50-minute sessions over 2 weeks.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Darch &amp; Camine (1986)</strong></td>
<td>• T (Visual display): A visual display of the concept relationships within each unit.</td>
<td>T vs. C (posttest): ES = 1.34, p &lt; .01</td>
<td>T vs. C (follow-up): ES = 0.36, ns</td>
</tr>
<tr>
<td>24 fourth, fifth, and sixth graders with LD</td>
<td>• C (Text): Typical teacher-initiated group instruction.</td>
<td>T vs. C (transfer test): ES = −0.67, ns</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Duration/Intensity: Nine 50-minute sessions over 2 weeks.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Darch &amp; Eaves (1986)</strong></td>
<td>• T (Visual display): Same as Darch &amp; Camine (1986)</td>
<td>T vs. C (posttest): ES = 1.78, p &lt; .01</td>
<td></td>
</tr>
<tr>
<td>22 high school students with LD</td>
<td>• C (Text): Same as Darch &amp; Camine (1986)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Duration/Intensity: Twelve 50-minute sessions over 3 weeks.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Darch &amp; Gersten (1986)</strong></td>
<td>• T (Advanced organizer): An outline or overview of facts and concepts from a lesson that showed the relationships among the concepts.</td>
<td>T vs. C (posttest): ES = 0.47, ns</td>
<td></td>
</tr>
<tr>
<td>24 high school students with LD</td>
<td>• C (Basal reading): Lesson through a lecture and discussion format with students’ independent reading and self-study.</td>
<td>T vs. C (posttest): ES = 0.43, ns</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Duration/Intensity: Nine 50-minute sessions over 2 weeks.</td>
<td>T vs. C (follow-up): ES = −0.16, ns</td>
<td></td>
</tr>
<tr>
<td><strong>Griffin, Simmons, &amp; Kame’enui (1991)</strong></td>
<td>• T (Graphic organizers): Graphic organizers reflecting the hierarchical relationships among the ideas in the chapters.</td>
<td>T vs. C (posttest): ES = 0.59, ns</td>
<td></td>
</tr>
<tr>
<td>28 fifth and sixth graders with LD</td>
<td>• C (List of facts): Vertical lists of information.</td>
<td>T vs. C (follow-up): ES = 0.32, ns</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Duration/Intensity: Four 45-minute sessions over 1 week.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. T = treatment; C = comparison; ES = effect size; ns = nonsignificant; LD = learning disabilities; EMR = educable mental retardation; EMH = educable mental handicapped; CB = curriculum-based measure; SDRT = Stanford Diagnostic Reading Test.

*ES was reported in an original study, but the study did not provide results from significant tests.

---

**Study** (Sinatra, Stahl-Gemake, & Berg, 1984), whereas Horton et al. (1990), Studies 1, 2, & 3 examined both junior high and high school students in three of their studies.

**Research Design**

All of the studies employed group designs. More specifically, studies employed either a treatment–comparison design (n = 16) or a single-group with multiple treatments design (n = 5).

**Intervention Variables**

**Types of Interventions.** Although all of the studies used graphic organizers in their interventions, researchers chose a variety of graphic organizers. A cross-study analysis revealed four types of graphic organizers:

1. **Semantic organizers** represent relationships between concepts and features of concepts (e.g., related vocabulary) and include semantic mapping, semantic feature analysis, and semantic/syntactic feature analysis (n = 9; Anders et al., 1984; Bos & Anders, 1990; Bos & Anders, 1992, Studies 1, 2, 3, 4, 5, & 6; Bos et al., 1985, 1989).

2. **Cognitive maps with a mnemonic** are implemented by a teacher/researcher or student to construct a cognitive map through the use of a mnemonic (e.g., TRAVEL, RELATE) that displays concept relationships in a unit (n = 3; Boyle, 1996, 2000; Boyle & Weishaar, 1997).

3. **Cognitive maps without a mnemonic** display the various concept relation-
4. Framed outlines consist of graphically represented lesson outlines or overviews that help teachers and students identify main ideas and important facts (n = 2; Darch & Gersten, 1986; Lovitt et al., 1986).

Duration/Intensity of Interventions. In 19 of the studies, the interventions lasted between 1 week and 3 weeks, resulting in a range of 2 to 12 sessions. The interventions in the other two studies lasted 12 to 16 weeks with an unreported total number of sessions (Lovitt et al., 1986; Sinatra et al., 1984).

Persons Implementing Interventions. Across the studies, either teachers (n = 13), researchers (n = 6), or...
teachers–researcher (n = 2) delivered the interventions.

**Persons Generating Graphic Organizers.** Teachers and researchers generated the graphic organizers in 16 of the studies. The graphic organizers in these studies included two types:

1. those generated prior to instruction and used during instruction (e.g., Darch & Gersten, 1986; Horton et al., 1990, Study 1), and
2. those generated prior to instruction but left incomplete for the teacher or researcher to fill out with the students during instruction (e.g., Anders et al., 1984).

In four other studies, the students themselves generated one of two different types of graphic organizers:

1. independently student-generated ones (Boyle, 1996, 2000), or
2. blank ones designed by the teacher and filled in by the students independently (Horton et al., 1990, Studies 2 & 3).

One final study included both teacher-generated and student-generated graphic organizers and compared their effects (Boyle & Weishaar, 1997).

**Comprehension Measures**

Across the studies, several different types of measures were used to assess reading comprehension. In a majority of the studies (n = 18), only researcher-developed tests were used. Two studies included both researcher-developed tests and standardized reading tests (Boyle, 1996; Boyle & Weishaar, 1997), and one study used a curriculum-based assessment developed by a textbook publishing company (Lovitt et al., 1986).

**Findings**

We summarized the findings from intervention studies by examining the effects by type of graphic organizers, grade levels of participants (i.e., elementary, junior high, or high school), persons implementing interventions, persons generating graphic organizers, types of measurements, and additional measures (i.e., follow-up tests and transfer tests). We reviewed treatment–comparison design studies first, and then single-group design studies were reviewed to confirm or disconfirm these initial findings.

**Types of Graphic Organizers.** The use of different graphic organizers (i.e., semantic organizers, cognitive maps with and without a mnemonic, and framed outlines) and the effect sizes associated with the use of each organizer are presented.

**Semantic organizers.** Nine treatment–comparison design studies by Bos and colleagues revealed consistent findings for the effects of semantic organizers on reading comprehension. Students with LD who used semantic organizers demonstrated significantly higher scores on researcher-developed comprehension measures than students in comparison groups (i.e., receiving dictionary instruction, normative comparison group). In eight of the nine studies, semantic organizers were associated with large effect sizes (d = 0.81–1.69). Semantic organizers in the ninth study were associated with a small to medium effect size (d = 0.40; Bos & Anders, 1992, Study 5). The relatively small effect size in this study (Bos & Anders, 1992, Study 5), however, may have been caused by the comparison group being composed of students without LD and actually represents a more positive outcome—students with LD using semantic organizers performed slightly better than typically developing students.

**Cognitive maps without a mnemonic.** Seven studies examined the effects of cognitive maps without a mnemonic on students’ reading comprehension (Darch & Carnine, 1986; Darch & Eaves, 1986; Griffin et al., 1991; Horton et al., 1990, Studies 1, 2, & 3; Sinatra et al., 1984). The use of cognitive maps was associated with higher comprehension scores than comparison conditions (i.e., typical reading instruction) in two treatment–comparison group design studies (d = 1.79 and 1.34 respectively; Darch & Carnine, 1986; Darch & Eaves, 1986). Four single-group design studies also demonstrated the positive effects of using cognitive maps on students’ reading comprehension.
hension ($d = 0.96$-$5.07$; Horton et al., 1990, Studies 1, 2, & 3; Sinatra et al., 1984). On the other hand, Griffin et al. (1991) found that such cognitive maps did not significantly enhance the reading comprehension of students with LD ($d = 0.50$, $p > .05$); however, this nonsignificant finding may have resulted from the comparison condition in this study being very similar to the intervention condition. For instance, both groups received similar instruction (i.e., a structured overview) in the Griffin et al. (1991) study, whereas in other studies, students in the comparison groups received different instruction (e.g., using dictionaries, reading basals) from those in the intervention groups using cognitive maps.

Framed outlines. Two studies examined the effects of using framed outlines on the reading comprehension of students with LD. One treatment–comparison group design study found that students using framed outlines significantly outperformed those in comparison conditions (i.e., reading basals) on a reading comprehension test ($d = 1.78$; Darch & Gersten, 1986). This finding was confirmed by a single-group design study that demonstrated positive effects of using framed outlines on students' reading comprehension ($d = 0.80$; Lovitt et al., 1986).

Participants' Grade Levels. For this particular analysis, we included only those studies conducted at one school level (i.e., elementary, junior high, or high school). We found no significant differences across grade levels for the effects of using graphic organizers on students' reading comprehension scores. The use of graphic organizers consistently related to large effect sizes across grade levels. For example, high school students using graphic organizers received higher comprehension scores than those in comparison conditions (e.g., using dictionaries, self-study; $d = 1.18$; $k = 10$). For junior high school students, graphic organizer use was associated with a mean ES of 1.08 ($k = 13$) in treatment–comparison design studies and with a mean ES of 0.80 ($k = 1$) in a single-group design study. Finally, graphic organizers were associated with a mean ES of 0.79 ($k = 7$) for elementary school students in treatment–comparison design studies.

Persons Implementing Interventions. Whether researchers, teachers, or researcher–teachers implemented the intervention, the use of graphic organizers was associated with large effect sizes in treatment–comparison design studies ($d = 0.96$, $k = 17$; $d = 1.05$, $k = 12$; and $d = 1.61$, $k = 2$; respectively). A mean ES of 2.38 ($k = 5$) was also found for the single-group design studies in which a teacher led the graphic organizer instruction.

Persons Generating Graphic Organizers. Regardless of who generated the graphic organizers (i.e., teachers, researchers, or students), large effect sizes were found. For studies in which teachers or researchers generated the graphic organizers, the mean effect sizes ranged from 1.15 ($k = 19$) for treatment–comparison design studies to 1.20 ($k = 3$) for single-group design studies. Similarly, large effect sizes were found for studies using student-generated graphic organizers: 0.86 ($k = 12$) for treatment–comparison design studies and 4.15 ($k = 2$) for single-group design studies. One study compared the effects of using a student-generated graphic organizer with those of using a researcher-generated graphic organizer (Boyle & Weishaar, 1997). The results revealed that the use of a student-generated graphic organizer yielded higher comprehension scores than did the implementation of a researcher-generated graphic organizer on one of four comprehension measures ($p < .05$). Two additional single-group design studies also demonstrated that using student-generated graphic organizers was associated with larger effect sizes than using teacher-generated graphic organizers (Horton et al., 1990, Studies 1 & 2). It is important to note, however, that regardless of who generated the graphic organizers (i.e., teachers, researchers, or students), groups using such visual displays significantly outperformed those in comparison conditions (i.e., conventional reading techniques, self-study) in these two studies (Horton et al., 1990, Studies 1 & 2).

Types of Measures. Twenty of the 21 studies used researcher-developed comprehension tests as their primary dependent measures, and only 2 of these 20 studies also included standardized reading tests (Boyle, 1996; Boyle & Weishaar, 1997). Intervention effects based on researcher-developed reading tests were significantly larger than those based on standardized tests. In fact, neither of the two studies using standardized reading tests reported significant differences between the graphic organizer intervention (i.e., cognitive maps with mnemonics) and comparison conditions (Boyle, 1996; Boyle & Weishaar, 1997). However, when researcher-developed tests were used as the comprehension measures, students using cognitive maps performed significantly better than students in the comparison group (Boyle, 1996; Boyle & Weishaar, 1997).

Additional Measures. In addition to posttest comprehension measures, eight of the treatment–comparison design studies conducted follow-up tests that contained questions similar to those on the posttests about the reading passage used in the intervention. Although the differences between the graphic organizer and comparison groups were greater on posttests ($d = 1.04$) than on follow-up tests ($d = 0.60$), the mean effect size obtained on follow-up tests was still large.

Two treatment–comparison design studies included transfer tests, which assessed students’ abilities in applying a graphic organizer to a new reading passage (Darch & Carnine, 1986; Darch & Eaves, 1986). No statistically significant differences between graphic organizer groups and comparison groups
on transfer tests were found for either study.

**Discussion**

Our examination of the effects of graphic organizers on the reading comprehension of students with LD revealed overall beneficial outcomes across the studies. In general, our findings support the use of semantic organizers, cognitive maps with and without mnemonics, and framed outlines to promote students’ reading comprehension. Across the board, when the students were taught to use graphic organizers, large effect sizes were demonstrated on researcher-developed reading comprehension posttests. Thus, visual displays of information such as those provided by graphic organizers enhance the reading comprehension of students with LD, possibly by helping these students organize the verbal information and thereby improving their recall of it.

The results reported in this synthesis are impressive given that a majority of the 21 studies reviewed included both treatment and comparison conditions. The inclusion of comparison conditions makes it possible for significantly higher effects to be due to the intervention (i.e., using graphic organizers) rather than to general learning or factors other than the intervention. The changes in comprehension reported in these treatment–comparison design studies, therefore, are greater than what would be expected from standard instruction.

Of the three studies reviewed that did not yield significantly more positive effects for the intervention group than for the comparison group, two studies (Bos & Anders, 1992, Studies 5 & 6) included average achievers in their comparison groups instead of students with LD. In these two studies, the researchers trained the intervention group (i.e., students with LD) to use graphic organizers and then compared their performance to that of average-achieving students who did not receive such training—a comparison that yielded small to medium effect sizes. Due to the differences in the general academic achievement of these two groups, it is difficult to interpret these results. A better test of the effectiveness of using graphic organizers would be to assign students with LD randomly to either the graphic organizer or the comparison condition and not to include other types of students.

The third study (Griffin et al., 1991) in which nonsignificant results were found included a comparison group who received similar instruction (i.e., a structured overview) to the intervention group but in a different format. The intervention group received the information in the form of a graphic organizer, whereas the comparison group was given a list format. It is important to note that this was the only study reviewed in which the comparison group received the same key information from the text as the intervention group. Based on this study’s results, it may be the practice of extracting important information from the text that promotes reading comprehension, and not the specific graphic organizer per se; however, further research on specific graphic organizer features is needed to support this interpretation.

With respect to the effectiveness of using graphic organizers across grade levels, our analysis demonstrated effective outcomes at the elementary and secondary levels for the use of graphic organizers. Only one study (Sinatra et al., 1984) included students at the lower elementary level (i.e., second and third grades), and the researchers did not disaggregate the data for this sample of students from that of older elementary students, thereby making a separate interpretation of the benefits for lower elementary students difficult. Due to the lack of focus that this and other studies placed on lower elementary school students, the positive outcomes demonstrated for the use of graphic organizers in our analysis are limited to the upper elementary and secondary grade levels.

In addition to analyzing the effects by grade levels, we examined the influence that those implementing the intervention had on the outcomes. Both teacher- and researcher-directed graphic organizer interventions were associated with large effect sizes. In fact, higher effect sizes resulted from teachers rather than researchers leading the graphic organizer instruction in treatment–comparison design studies ($d = 0.96$ and $1.05$ respectively)—a finding that is slightly different from a previous synthesis that examined the effects of multiple interventions on students with LD and revealed that researcher-directed interventions were generally associated with more positive outcomes than teacher-directed ones (Swanson et al., 1999). Rather than having a researcher or a teacher lead the intervention, a few of the studies included what they called a teacher–researcher as the director of the intervention (Anders et al., 1984; Bos et al., 1985, 1989). Although the interventions under the teacher–researcher’s direction resulted in higher effect sizes than the teacher-directed interventions, these studies did not provide a clear definition of what they meant by teacher–researcher, which makes the difference between these studies and the other ones difficult to understand. Nevertheless, the high effect sizes obtained from the teacher-directed interventions show the feasibility of implementing graphic organizer techniques in the classroom setting.

Although all of these different analyses revealed positive effects of using graphic organizers to improve the reading comprehension of students with LD, it is important to note that the reading comprehension measures that a majority of these studies used were developed by the researchers themselves and were aligned closely with the intervention content. In fact, only two studies included standardized reading tests, and neither of these studies reported statistically significant differences between the intervention and comparison conditions when such tests were used. Furthermore, despite...
the findings that students participating in the interventions outperformed those in comparison groups on reading comprehension assessments, few participants reached comprehension levels above 80% on the measures. In studies yielding significant effect sizes, the mean posttest scores ranged from 48% to 83% correct, with 15 of the 29 posttests yielding means below 70%. Based on these findings, then, the use of graphic organizers is not sufficient for ensuring that students will obtain acceptable levels of understanding.

Related to this finding, high effect sizes were found on posttests and follow-up tests but not on transfer tests. In a majority of the studies using follow-up tests, both these and the posttests addressed the specific passages that students had been taught previously, which means essentially that the use of graphic organizers was an effective instructional tool for teaching the content of specific passages. However, these studies provided no evidence to show that students could transfer the use of graphic organizers to understand other passages. In fact, no statistically significant differences were found when researchers used transfer assessments—meaning that students did not transfer their skills (i.e., using graphic organizers) to novel reading situations. These findings are supported by similar results reported in another synthesis focusing on reading comprehension strategies in general (Gersten, Fuchs, Williams, & Baker, 2001). One explanation for the findings in our synthesis might be these studies’ short intervention times (i.e., 1–3 weeks in 19 studies). Longer intervention durations may be necessary for students to use graphic organizers more easily and independently in novel situations. To support this hypothesis, the use of graphic organizers related to more positive outcomes when the intervention offered longer instruction and training (Alvermann & Boothby, 1986; Bean, Singer, Sorter, & Frazee, 1986). Along with longer interventions, researchers may need to include specific instruction on how to generalize the use of graphic organizers to new reading situations—a hypothesis supported by other research showing that students with LD often have difficulty learning to generalize without specific instruction (Deshler & Schumaker, 1993).

Methodological Issues

Findings from this synthesis confirmed the findings of previous reviews on graphic organizers used with students without disabilities showing that there has been a lack of consistent operational criteria for what constitutes a graphic organizer. In our corpus of studies, the graphic organizers varied by type (e.g., semantic organizers, framed outlines, originators (i.e., teacher, researcher, or student), and person implementing the intervention (Dunston, 1992; Moore & Readence, 1980, 1984). However, the duration of the interventions and outcome measures were fairly consistent across the studies; a majority of them lasted for a short time and assessed students’ reading comprehension with researcher-developed posttests.

Whereas our overall findings are positive, several methodological issues in the studies limit the practical implications of these results. First, the type of measures used to assess comprehension performance was one major concern of the study. All of the studies yielding significant results used researcher-developed assessments, whereas none of the studies using standardized reading tests reported significant results on such tests. Paralleling these findings, previous meta-analyses revealed that higher effect sizes were more likely to be obtained on nonstandardized measures developed by researchers than on standardized tests (Elbaum, Vaughn, Hughes, & Moody, 2000; Swanson et al., 1999). Furthermore, most of the studies used the same passage for both the intervention’s implementation and its assessment. Taking our findings in conjunction with these meta-analyses’ results, we can conclude that participants in our synthesis made gains in the comprehension of the content taught, but whether such gains would generalize to other independent reading situations and what part the graphic organizers played in this process are not clear.

A second issue relates to the comparison conditions used in the treatment–comparison design studies and whether they were robust enough to provide high standards against which the effects of graphic organizer interventions could be measured (Gersten, Baker, & Lloyd, 2000; Pressley & Harris, 1994). Only one study compared the use of graphic organizers to other specific comprehension strategies (i.e., a structured overview) that might be considered comparable with the intervention condition. In all of the other studies, the researchers compared their interventions with much more distal forms of instruction (e.g., typical reading instruction) and, thereby, demonstrated only that using graphic organizers is superior to typical instruction in reading or content-area comprehension. Research comparing graphic organizer interventions to other, more specific reading comprehension strategies would help educators to ascertain whether using graphic organizers is a less strong, equally strong, or superior strategy for comprehending text.

A third issue relates to the duration of the interventions. The duration of most of the interventions was short, ranging between 1 week and 3 weeks, with only two interventions lasting longer than 10 weeks (Lovitt et al., 1986; Sinatra et al., 1984; 12 weeks and 16 weeks, respectively). Such short intervention periods may relate to the limited effects found for the use of graphic organizers on more generalized measures such as standardized reading tests and transfer tests.

A fourth issue is that a majority of the interventions used graphic organizers generated by teachers or researchers, instead of teaching students to generate their own graphic organizers. Only four studies implemented the use of graphic organizers as a student-directed strategy. Rather than focusing
on how graphic organizers act as learning tools to help students become more independent readers, most of the studies in this synthesis focused on how graphic organizers can act as instructional tools for teachers to improve their students’ reading comprehension of specific texts.

Finally, the same research teams conducted the research on particular types of graphic organizers (e.g., semantic organizers by Bos and colleagues; cognitive maps using mnemonics by Boyle and colleagues). Such a lack of corroboration across research teams and their interventions limits the reported evidence about the use of graphic organizers. Replication of the interventions by other researchers would help to ascertain whether the use of different types of graphic organizers is an effective strategy for improving students’ reading comprehension.

**Limitations**

In this synthesis, we included the findings from all of the group design studies conducted between 1963 and June 2001 that examined how effective the use of graphic organizers is at improving the reading comprehension of students with LD. Our synthesis did not, however, include single-subject design studies. Although we recognize the importance of examining such studies, previous syntheses of interventions have reported separate summaries for group design and single-subject design studies (e.g., Swanson et al., 1999), and we chose to model our synthesis after these previous syntheses. Future research to address single-subject design studies on graphic organizers is warranted to extend this current finding. Despite the lack of a formal search for and analysis of such studies, we did include them during our hand search of 10 journals to make an informal comparison between this study’s findings and our synthesis’ findings. We uncovered one article examining the effects of using story maps on the reading comprehension of students with LD (Gardill & Jitendra, 1999). Interestingly, this study did not focus on using such graphic organizers to aid students with the comprehension of expository text, like the studies included in our synthesis, but instead related this strategy to understanding narrative text.

**Implications for Future Research**

Although the current literature provides evidence supporting the use of graphic organizers to promote reading comprehension among students with LD, future research on such graphic organizer use should focus on several factors including measuring comprehension with standardized reading tests, examining interventions with lower elementary school students, using graphic organizers with independent reading, comparing the use of complex versus simple graphic organizers, comparing graphic organizer interventions with other comprehension strategies, and replicating the present research findings.

As described in our results, the first of these factors—measuring reading comprehension with standardized reading assessments—was not addressed in a majority of studies included in our synthesis. Furthermore, the two studies that included standardized reading tests did not demonstrate positive effects for the use of graphic organizers on such tests. Thus, further research examining the effects of using graphic organizers on reading comprehension as measured by standardized tests is warranted to ascertain whether graphic organizers are effective tools for improving scores on more distal comprehension measures.

A second area for future research focuses on the types of students who can benefit from using graphic organizers. The reviewed studies conducted graphic organizer interventions with upper elementary, junior high, and high school students, which demonstrates a lack of information about how to use graphic organizers with younger students (i.e., lower elementary students). Related to this, it is unknown if students need to read at a certain level before being able to use graphic organizers effectively. If this is the case, future lines of research might include examining the effects of using graphic organizers with students at various reading levels and on younger students’ listening comprehension (i.e., a possible precursor to reading comprehension).

Third, as discussed earlier, the literature has focused more on improving students’ reading comprehension through teacher instruction with graphic organizers than through students’ independent use of graphic organizers. Further research needs to clarify whether students’ use of graphic organizers independently is an effective strategy to enhance their reading comprehension. Researchers must also keep in mind both the type of graphic organizers (e.g., semantic organizers, framed outlines) and the duration of instruction as factors influencing how effective students’ use of graphic organizers independently can be at improving students’ reading comprehension.

A fourth potentially important research area involves examining the effectiveness-efficiency ratio for specific types of graphic organizers. Specifically, researchers must further investigate the differential effects of using complex versus simple graphic organizers on students’ comprehension and the differences in teacher sustainability of using complex versus simple graphic organizers in their instruction. For example, some graphic organizers (i.e., semantic feature analysis) take a long time to teach from or complete, and if student gains in comprehension are minimal, the use of such organizers may be of little value.

A fifth area on which future research might focus lies in comparing the effects of using graphic organizers with other comprehension strategies. Knowledge of the conditions under which graphic organizer use is superior or inferior to other reading strategies will assist educators in choosing the appro-
appropriate comprehension instruction for students with LD. Furthermore, studies using these types of comparisons and demonstrating positive effects for using graphic organizers will generate greater confidence in such interventions as helpful instructional practices for teaching students with LD.

In addition to more in-depth studies of graphic organizer use within these different areas, future research needs to replicate the specific graphic organizer methods explored in the present studies. Only three studies examining the use of graphic organizers with students with LD have been conducted in the last 10 years. Moreover, generally speaking, each research team studied one type of graphic organizer. Replication of these studies by various researchers across different settings is necessary in order to generalize about the benefits of the intervention techniques.

ABOUT THE AUTHORS

Ae-Hwa Kim, PhD, is a member of the faculty in the Dept. of Special Education at Dankook University in Seoul, Korea. Her research interests include effective reading interventions and technology applications for students with learning disabilities. Sharon Vaughn, PhD, is the H. E. Hartfelder/Southland Corp. Regents Professor and director of The University of Texas at Austin Center for Reading and Language Arts. Her research addresses social and academic outcomes for students with learning disabilities and reading difficulties. Jeanne Wanzek, MA, is a doctoral student in special education at The University of Texas at Austin. Her research interests include effective instructional practices for students with learning disabilities. Shang-jin Wei, MA, is a doctoral student in special education at The University of Texas at Austin. Address: Ae-Hwa Kim, Dankook University, Department of Special Education, San #8, Hannam-dong, Yongsan-gu, Seoul, Korea 140-714; aehwa@dankook.ac.kr

REFERENCES


one-to-one tutoring programs in reading for elementary students at risk for reading failure? Journal of Educational Psychology, 92, 605–619.


Copyright of Journal of Learning Disabilities is the property of PRO-ED and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.