Learning Disability Quarterly
2011, Vol. 34, No. 1, 51-72

GRAPHIC ORGANIZERS AND STUDENTS WITH LEARNING DISABILITIES: A META-ANALYSIS

Douglas D. Dexter and Charles A. Hughes

Abstract. This meta-analysis reviews experimental and quasi-experimental studies in which upper-elementary, intermediate, and secondary students with learning disabilities learned from graphic organizers. Following an exhaustive search for studies meeting specified design criteria, 55 standardized mean effect sizes were extracted from 16 articles involving 808 participants. Students at levels ranging from grade 4 to grade 12 used graphic organizers to learn in core-content classes (English/reading, science, social studies, mathematics). Posttests measured near and far transfer. Across several conditions, settings, and features, the use of graphic organizers was associated with increases in vocabulary knowledge, comprehension, and inferential knowledge. Mean effect sizes varied from moderate to large based on type of measure, type of graphic organizer, and subject area. Conclusions, implications for future research, and practical recommendations are presented.

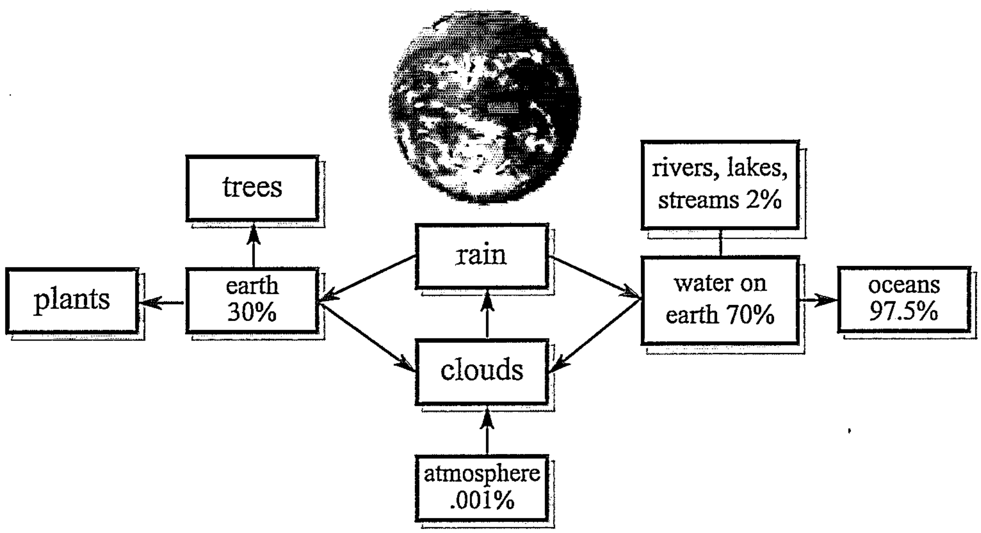
DOUGLAS D. DEXTER, Ph.D., College of Education, The Pennsylvania State University, University Park, Pennsylvania.
CHARLES A. HUGHES, Ph.D., College of Education, The Pennsylvania State University, University Park, Pennsylvania.

As students enter the upper-elementary, intermediate, and secondary grades, academic demands are heightened as material becomes more complex and the curriculum is driven by higher-order skills and advanced concepts (Fletcher, Lyon, Fuchs, & Barnes, 2007). Students at these levels typically receive less individual attention than in primary grades (Hughes, Maccini, & Gagnon, 2003), and are often required to learn primarily through didactic lecture and expository text presentation (Gajria, Jitendra, Sood, & Sacks, 2007; Minskoff & Allsopp, 2003). This shift in learning presentation, rife with abstract concepts, unfamiliar content, and technical vocabulary (Armbruster, 1984), may seem daunting to most students, but especially so to students with learning disabilities (LD).

Students with LD often have difficulty with basic academic skills (e.g., reading) and organizational/study skills (Deshler, Ellis, & Lenz, 1996). These students generally have difficulty connecting new material to prior knowledge, identifying and ignoring extraneous information, identifying main ideas and supporting details, drawing inferences, and creating efficient problem-solving strategies (Baumann, 1984; Holmes, 1985; Johnson, Graham, & Harris, 1997; Kim, Vaughn, Wanzek, & Wei, 2004; Williams, 1993). Because many textbooks are written above grade-level reading ability and lack organizational clarity (Gajria et al., 2007), these learning difficulties make interpreting and comprehending expository text especially challenging (Bryant, Ugel, Thompson, & Hamff, 1999).

Students with LD need explicit content enhancements to assist in verbal (e.g., text or lecture) comprehension, and graphic organizers (GOs) have often been recommended as an instructional device to assist these

Figure 1. Cognitive mapping example (science).



students in understanding increasingly abstract concepts (Bos & Vaughn, 2002; Hughes et al., 2003; Ives & Hoy, 2003; Kim et al., 2004; Nesbit & Adesope, 2007; Rivera & Smith, 1997).

What Are Graphic Organizers?

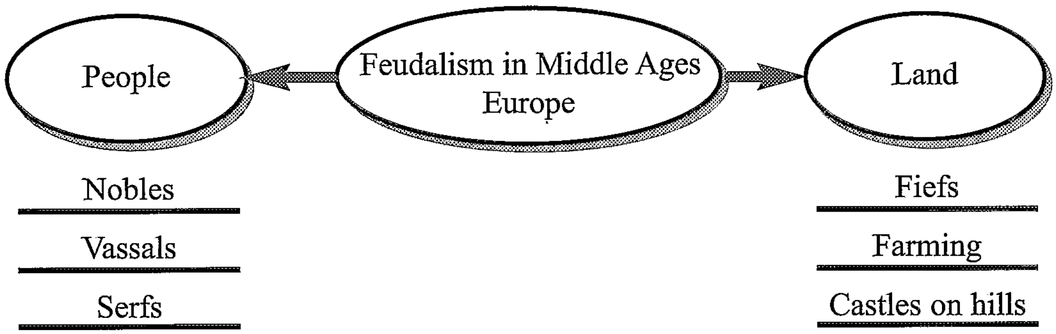
GOs are visual and spatial displays that make relationships between related facts and concepts more apparent (Gajria et al., 2007; Hughes et al., 2003; Kim et al., 2004). They are intended to promote more meaningful learning and facilitate understanding and retention of new material by making abstract concepts more concrete and connecting new information with prior knowledge (Ausubel, 1968; Mayer, 1979). While there is inconsistency in the definitions of types of GOs (Rice, 1994), we were able to classify all the studies on students with LD where effects could solely be attributed to the GO into the following five general categories.

Cognitive Mapping. Cognitive mapping assists in making major ideas and relationships explicit by using "lines, arrows, and spatial arrangements to describe text content, structure, and key conceptual relationships" (Darch & Eaves, 1996, p. 310). By graphically displaying important information, cognitive mapping may support and develop organizational strategies essential to both reading and writing of content area material by bridging the gap between idea organization and written language structures (Pehrssen & Denner, 1988). Further, Boyle and Yeager (1997) pointed out that minimizing sentences and details in the GO is an essential component of cognitive mapping. They recommend using keywords and simple drawings rather than complex sentences or elaborate drawings. Teachers often make cognitive mapping GOs to use as advance organizers for challenging text. Students also can fill in teacher-prepared blank cognitive mapping GOs during or after attending to challenging verbal material. An example of a cognitive mapping GO is found in Figure 1.

Semantic Mapping. Semantic mapping (SM) is a heuristic that enables students to recognize relevant information from lecture and text (e.g., main ideas, important supporting details), delete isolated details that may not be relevant to overall understanding, and highlight key concepts that may have not been fully developed in a lecture or text (Bos & Anders, 1990).

Unlike cognitive mapping, when using SM, students and the teacher actively create a visual representation (e.g., relationship map or web) to represent the rela-

Figure 2. Semantic mapping example (social studies).



tionships among concepts. Typically, concepts are listed, and the teacher and students make predictions about how the concepts can be arranged to demonstrate those relationships on the GO (Bos & Anders, 1992).

A well-made GO consists of a superordinate concept (e.g., main idea, topic) placed in an oval in the middle or top of the page. Coordinate concepts (e.g., categories representing related concepts) are placed in ovals surrounding or underneath the superordinate concept and connected by lines. Coordinate concepts can include a variety examples, functions, or characteristics of the superordinate concept. Finally, subordinate concepts (e.g., concepts representing the coordinate concept) are listed below each coordinate concept (Bos & Anders, 1990, 1992; Pearson & Johnson, 1978). Figure 2 provides an example of an SM GO.

Semantic Feature Analysis. Semantic feature analysis (SFA) is similar to SM by helping students recognize relevant information from lecture and text. This is done through a presentation of related concept characteristics in a matrix form. In SFA, unrelated concepts can be inferred directly from the chart (Darch &

Figure 3. Semantic feature analysis example (science).

Comparison of Dog Breeds

	Basset Hound	**Old English Sheepdog**	**Brittany Spaniel**	**Border Collie**
Energetic	–	+	+	+
Weekly Grooming	–	+	–	–
Good with Kids/Other Pets	+	+	?	–

Note. + = feature present; - = feature not present; ? = not sure.

Gersten, 1986). Typically, a relationship matrix is constructed with vocabulary representing the coordinate concepts placed along the top of the matrix, and the vocabulary representing the subordinate concepts placed along the side (Bos & Anders, 1990) The teacher and students can then make predictions and confirmations of relationships (e.g., related, not related, not sure) between the coordinate and subordinate concepts (Bos &Anders, 1992). The superordinate concept serves as the title. Figure 3 provides an example of an SFA G

Syntactic/Semantic Feature Analysis. Syntactic/semantic feature analysis (SSFA) is nearly identical to SFA but with the addition of cloze-type sentences written based on the matrix (Bos & Anders, 1990). Cloze sentences contain blank spaces replacing new vocabulary words. Students must use the context of the sentence and the SFA matrix to fill in the blanks. An example of an SSFA GO is found in Figure 4.

Visual Display. Visual displays present concepts or facts spatially, in a computationally efficient manner. That is, relationships between concepts are made apparent and clear by their location on the display. According to Hughes et al. (2003), in a visual display, facts or concepts are typically presented in one of five ways: temporal (e.g., timeline), spatial (e.g., decision tree), sequential (e.g., flowchart), hierarchal (e.g., taxonomy), or comparative (e.g., Venn diagram). An example of a visual display GO is found in Figure 5.

Previous Research

Several groups of researchers have conducted reviews and meta-analyses of the effectiveness of using GOs with nondisabled students (e.g., Ausubel, 1968; Kulhavy, Stock, & Caterino, 1994; Mayer, 1979; Moore & Readence, 1984; Nesbit & Adesope, 2007; Robinson, Katayama, DuBois, & Devaney, 1998). Based on these examinations of both the benefits of GOs and the effective design of GOs, several key findings are consistently replicated: (a) students with low verbal ability gain more from GOs than students with high verbal ability; (b) students with little or no prior knowledge in a subject gain more from GOs than students with an abundance of prior knowledge in a subject; (c) GOs are especially helpful in assisting students with far-transfer tasks, in addition to near-transfer tasks and factual recall; (d) GOs should be explicitly taught to students for maximum impact; (e) GOs should spatially group together or connect concepts so readers are more likely to perceive them as being interrelated and to draw

Figure 4. Syntactic/semantic feature analysis example (science).

Comparison of Dog Breeds

	Basset Hound	Old English Sheepdog	Brittany Spaniel	Border Collie
Energetic	–	+	+	+
Weekly Grooming	–	+	–	–
Good with Kids/ Other Pets	+	+	?	–

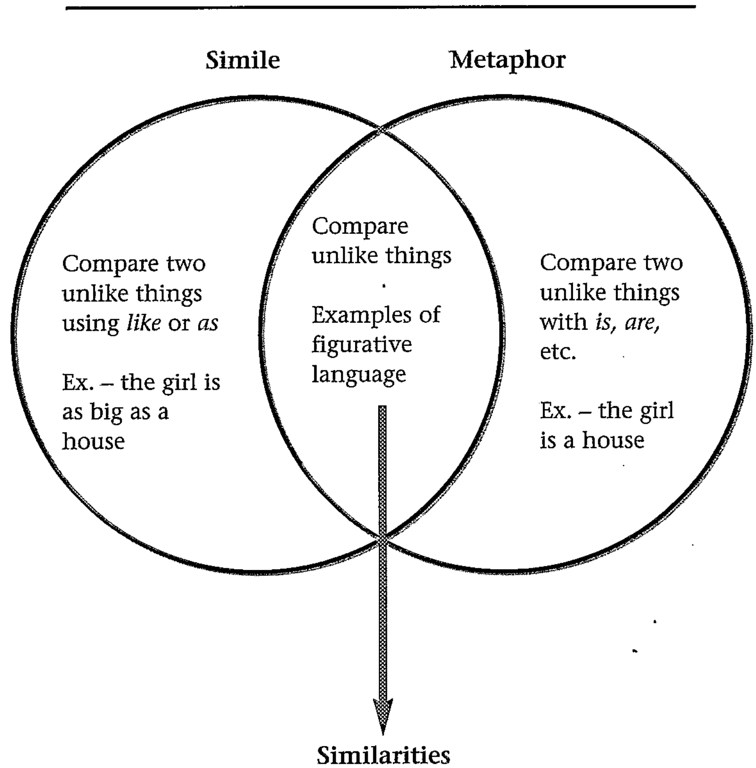
Note. + = feature present; - = feature not present; ? = not sure.

A(n) ______________ is the least energetic breed we have discussed.

A(n) ______________ requires at least weekly grooming.

A(n) ______________ and ______________ are good with other pets.

Figure 5. Visual dislay example (English/language arts).



perceptual inferences about their relationships; (f) GOs should not be clustered with a lot of information; readers should easily perceive the phenomena or relations that are important; (g) GOs are effective because of their computational efficiency, minimizing stress on the working memory; and (h) GOs can be effective when used before, during, or after a lesson. While these findings are promising, the vast majority of the studies reviewed used college students as participants, and few comparisons were made to students identified as LD.

Two research syntheses of school-aged children with LD (Gajria et al., 2007; Kim et al., 2004) have focused on GOs. In Gajria et al., as part of an examination of several content enhancements, GOs were found to have large effects for comprehension of expository text for upper-elementary, intermediate, and secondary students with LD. Likewise, Kim et al. found large effects for GOs on the reading comprehension of upper-elementary, intermediate, and secondary students with LD.

Drawing definitive conclusions from these reviews is problematic for several reasons, however. First, the reviews did not take sample sizes into consideration when calculating effect sizes and comparing studies. Estimations of effect were based on individual study characteristics without standardization. The use of meta-analysis allows for statistical standardization of findings, resulting in numerical values that are interpretable in a consistent pattern across all studies, thus controlling for an individual study's sample size (Lipsey & Wilson, 2001). Furthermore, according to Lipsey and Wilson, meaningful effects and relationships where multiple quantitative studies with varying sample sizes agree, as well as differential effects of study differences, are more likely to be identified by meta-analytic procedures than by less systematic and analytic approaches.

Second, the reviews focused solely on factual comprehension measures, not on vocabulary, inference, or relational comprehension. While information on factual comprehension is important, it may be useful to measure the utility of GOs for students with LD on multiple constructs rather than only on the ability to answer fact-level questions about a topic. The utility of an intervention (e.g., GO) on multiple constructs is considered important for high-quality research (Gersten, Baker, & Lloyd, 2000). Thus, the effectiveness of an intervention should be examined on all possible constructs.

Third, no systematic analysis of effects was conducted by type of measure (e.g., near or far transfer), type of GO, subject area, or student stage of attending to verbal material. This lack of clarity makes it difficult to identify why the GO interventions were so effective. A more systematic analysis of GO studies may allow for more precise explanations of effect based on multiple criteria, and lead to a more robust understanding of how and why GOs are effective for students with LD.

In the present study, we conducted a meta-analysis of GO research to address the following questions:

1. What are the overall effects of GOs on the posttest performance of students with LD?
2. Do these effects maintain over time?
3. Are there differential effects by type of measure (near or far transfer)?
4. Are there differential effects by type of GO?
5. Are there differential effects by subject area?
6. Are there differential effects by stage of attending to verbal material (before, during, after instruction)?

Method

Literature Search Procedure

A three-step process was used to identify studies using GOs with upper-elementary, intermediate, and secondary students with LD. First, we conducted a comprehensive computerized search of PsycInfo, ERIC, and Social Science Citation Index databases for studies from 1975 (e.g., year LD was officially recognized by passage of P.L. 94-142) to October 2009, using a list of search terms generated from previous reviews of GO studies (e.g., Horton et al., 1993; Moore & Readence, 1984; Nesbit & Adesope, 2006). We used the following combination of descriptors: *graphic organ*, expository, verbal, learning disab*, concept map, cognitive map, adolesc*, semantic map, semantic feature analysis, visual display.*

Second, we conducted ancestral searches of identified articles, as well as the two most recent reviews of content enhancements used with students with LD (e.g., Gajria et al., 2007; Kim et al., 2004).

Finally, we conducted hand searches of the following journals to locate the most recent literature: *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,* and *Reading Research Quarterly.* This process yielded a total of 27 published articles to analyze, many (e.g., 20) including more than one study.

Inclusion Criteria

We used six criteria to evaluate the appropriateness of each study. First, the study must have included a dependent measure of near or far transfer of verbal (e.g., text or lecture) material and a GO as the independent variable. Studies with a mnemonic illustration rather than GO as the content enhancement (e.g., Brigham, Scruggs, & Mastropieri, 1995; Mastropieri, Scruggs, & Levin, 1987) were excluded, as effects could not be attributed solely to the GO. Likewise, studies including GOs associated with Content Enhancement Routines (Bulgren, Deshler, Schumaker, & Lenz, 2000; Bulgren, Lenz, Schumaker, Deshler, & Marquis, 2002; Bulgren, Marquis, Lenz, Schumaker, & Deshler, 2009; Bulgren, Schumaker, & Deshler, 1988; Scanlon, Deshler, & Schumaker, 1996) were excluded, because effects could not be attributed solely to the GO, even though GOs were vital components of each of these routines. Further discussion about Content Enhance- ment Routines is included in the Limitations section of this meta-analysis.

Second, the study must have taken place in upper-elementary-, intermediate-, or secondary-level classrooms (e.g., grades 4-12). This grade range was selected because it represents a time when curricula typically become more complex and students are required to learn primarily through didactic lecture and expository text presentation (Fletcher et al., 2007; Hughes et al., 2003; Minskoff & Allsopp, 2003).

Third, based on the recommendations of Rosenthal (1994) and Lipsey and Wilson (1993), only studies using experimental or quasi-experimental group designs with control groups were included. Therefore, single-subject research studies (e.g., Gardill & Jitendra, 1999; Idol & Croll, 1987), repeated-measures studies (e.g., Lenz, Adams, Bulgren, Poulit, & Laraux, 2007), and single-group studies (e.g., Bergerud, Lovitt, & Horton, 1988; Boon, Fore III, Ayres, & Spencer, 2005; Horton, Lovitt, & Bergerud, 1990; Lovitt, Rudsit, Jenkins, Pious, & Benedetti, 1986; Sinatra, Stahl-Gemake, & Berg, 1984; Sturm & Rankin-Erickson, 2002) were excluded.

Fourth, the study must have provided sufficient quantitative information (e.g., group means and standard deviations; *F* statistic) to permit calculation of an effect size *(ES).* One experimental study (Boyle & Weishaar, 1997) was excluded because it provided only multivariate analysis of covariance (MANCOVA) data without group means and standard deviations. According to

Hunter and Schmidt (2004), there is no algebraically equivalent method to compute a comparable *ES* in such an instance.

Fifth, participants in the experimental and control groups must have included students with LD. We defined LD the same way as Kim et al. (2004) in their research review and Swanson, Hoskyn, and Lee (1999) in their meta-analysis (e.g., average intelligence and poor performance in at least one academic or related behavioral area). In each of the studies included, all participants were identified as students with LD.

Finally, the study must have been published in a peer-reviewed journal and in English. This excluded any studies in *Dissertation Abstracts International* and unpublished studies from researchers in the field. While this criterion ensures that only the highest quality research was included in the meta-analysis (Slavin, 1995), it also represents a potential publication bias (Lipsey & Wilson, 1993, 2001). This will be discussed further in the Limitations section.

Study Coding

The first author coded pertinent study features, including participant characteristics (i.e., grade level, disability classification), subject area, type of GO, stated purpose, study contrasts, dependent measures, and reported findings. A graduate research assistant double-coded this information, resulting in an interrater reliability of .97. After discussion and clarification to resolve disagreements in coding, interrater reliability reached 1.00.

Individual Effect Size Calculation

Using methods described by Lipsey and Wilson (2001), standardized mean difference effect size was computed using pooled standard deviation. The formula used was:

$$ES_{sm} = \frac{\overline{X}_{G1} - \overline{X}_{G2}}{S_p},$$

where "$\overline{X}_{G1}$ is the mean for group 1, $\overline{X}_{G2}$ is the mean for group 2, and S_p is the pooled standard deviation" (Lipsey & Wilson, p. 48). When studies only provided an *F* statistic, effect size was computed using a formula recommended by Thalheimer and Cook (2002). The formula used was:

$$d = \sqrt{F\left(\frac{n_t + n_c}{n_t n_c}\right)\left(\frac{n_t + n_c}{n_t + n_c - 2}\right)},$$

where n_t is the number of treatment subjects, and n_c is the number of control subjects.

Next, to correct for upwardly biased effect sizes due to small samples, a Hedges correction (Hedges, 1981; Lipsey & Wilson, 2001) was utilized. The unbiased effect size estimate was computed using the following formulae:

$$ES'_{sm} = \left[1 - \frac{3}{4N - 9}\right] ES_{sm},$$

$$SE_{sm} = \sqrt{\frac{n_{G1} + n_{G2}}{n_{G1} n_{G2}} + \frac{(ES'_{sm})^2}{2(n_{G1} + n_{G2})}},$$

$$w_{sm} = \frac{1}{SE^2_{sm}} = \frac{2 n_{G1} n_{G2} (n_{G1} + n_{G2})}{2(n_{G1} + n_{G2})^2 + n_{G1} n_{G2} (ES'_{sm})^2},$$

where *N* is the total sample size, ES_{sm} is the biased standardized mean difference, and

$$w_{sm} = \frac{1}{SE^2_{sm}}$$

is the inverse variance weight used to calculate the weighted mean effect size. According to Hedges, Shymansky, and Woodworth (1989), the inverse variance weight is a better approach to accounting for the sample size of a given study than the simpler approach of weighting by sample size.

Outliers

Prior to analyzing the weighted mean effect size, extreme effect sizes that may have disproportionate influence on the analysis were eliminated (Lipsey & Wilson, 2001). Based on the recommendation of Burns (2004), eliminated outliers were effect sizes that were greater than 1.5 times the mean effect size.

Data Analysis

Following transformations and outlier elimination, data were analyzed by computing the weighted mean effect size:

$$ES = \frac{\sum (w \times ES)}{\sum w}$$

where $\sum (w \times ES)$ is the summed product of the effect size and inverse variance weight, and $\sum w$ is the summed inverse variance weight.

Next, the standard error of the mean effect size was computed using the following formula:

$$se_{\overline{ES}} = \sqrt{\frac{1}{\sum w}},$$

where

$$\sqrt{\frac{1}{\sum w}},$$

is the square root of one divided by the summed inverse variance weight. The z-test for the weighted mean effect size was then computed by dividing the mean effect size by the standard error of the mean effect size, or in statistical notation:

$$Z = \frac{\overline{ES}}{se_{\overline{ES}}}$$

Finally, using the z-test, the 95% confidence interval for the weighted mean effect size was computed using the following formulae:

$$Lower = \overline{ES} - 1.96(se_{\overline{ES}})$$

$$Upper = \overline{ES} + 1.96(se_{\overline{ES}})$$

Homogeneity Analysis

Homogeneity analysis tests whether the assumption that all of the effect sizes estimate the same population mean is reasonable (Hunter & Schmidt, 2004). Lipsey and Wilson (2001) contend that single mean effect sizes by themselves are not sufficient descriptors of the distribution. Therefore, we computed a Q statistic to test homogeneity, using the formula:

$$Q = \sum (w \times ES^2) - \frac{[\sum (w \times ES)]^2}{\sum w}$$

The Q statistic is distributed as a chi-square with degrees of freedom (df) equaling number of *ESs* – 1. In our analysis of all studies, the critical value for a chi-square with df = 47 and $p = .05$ is 64. Because our calculated Q statistic (57.2) is less than this critical value, we can fail to reject the null hypothesis of homogeneity and assume a fixed-effects model, under which the variability across effect sizes does not exceed what would be expected based on sampling error. However, Lipsey and Wilson (2001) warned that a nonsignificant Q statistic "does not always provide great confidence that a fixed effects model is justified" (p. 117). Because we do not have a large number of effect sizes and the corresponding samples are relatively small, the Q statistic may not have sufficient statistical power. Therefore, we also fit a random-effects model, which assumes sampling error plus other sources of variability are randomly distributed (Lipsey & Wilson). The random-effects model is also a more conservative estimate than the fixed-effects model of differences between moderating variables (Hunter & Schmidt, 2004).

Whereas the fixed-effects model weights each study by the inverse of the sampling variance, the random-effects model weights each study by the inverse of the sampling variance plus a constant that represents the variability across the population effects (Lipsey & Wilson, 2001). The formulae are as follows:

$$w_i = \frac{1}{se_i^2 + \hat{v}_\theta},$$

$$\hat{v}_\theta = \frac{Q_r - k - 1}{\sum w - \left(\frac{\sum w^2}{\sum w}\right)},$$

where $\hat{v}_\theta$ is the random-effects variance component. We reran the analysis using this new weight to fit the random-effects model.

The preceding analyses were conducted for posttest measures, maintenance measures, and differential effects of individual levels of independent variables (e.g., GO type, subject area) and dependent variables (e.g., near- or far-transfer measures). Because it was possible to fail to reject the null hypothesis of homogeneity, we will present results for both the fixed-effects and random-effects model for comparison. Cohen's (1988) criteria for interpreting strength of effect sizes (small $ES < .20$, medium $ES = .50$, large $ES > .80$) were used to gauge the magnitude of the findings in this analysis.

Results

A total of 55 unique posttest effect sizes were extracted from studies in 16 published articles meeting our inclusion criteria. For the purposes of this analysis, each unique effect size was considered an individual estimate of effect. (Included articles are marked with an asterisk in the References section.) In addition, eight of the published articles included maintenance data rendering 29 more unique effect sizes. Table 1 includes detailed information on each study, participants, variables, measures, and individual effect sizes.

Instructional Context

Each of the studies included instruction on the use of a GO. The majority of studies (Anders, Bos, & Filip, 1984; Bos & Anders, 1990, 1992; Boyle, 1996, 2000; Darch & Carnine, 1986; Darch, Carnine, & Kame'enui, 1986; Darch & Eaves, 1986; Darch & Gersten, 1986; DiCecco & Gleason, 2002; Englert & Mariage, 1991; Griffin, Simmons, & Kame'enui, 1991; Hudson, 1996; Ives, 2007; Reyes, Gallego, Duran, & Scanlon, 1989) incorporated aspects of direct, explicit instruction (e.g., modeling, prompted practice). The authors of the remaining study (Bos, Anders, Filip, & Jaffe, 1989) reported that written guidelines for teaching the GO were developed; however, these guidelines were not included in the article.

Table 1

Independent Variables, Dependent Measures, and Effect Sizes of Individual Experiments

Subject/GO Type/Study/ Participants	Control Condition	Near or Far Transfer	Dependent Measure	Effect Size Posttest *M*	Maintenance *M*
SOCIAL STUDIES					
Semantic Mapping					
Bos & Anders (1992); Study 5: 26 upper-elementary students with LD	Normative	Near	Researcher-generated comprehension test	.42	.62*
	Normative	Near	Researcher-generated comprehension test	.30	.06
Semantic Feature Analysis					
Anders, Bos, & Filip (1984); 62 high school students with LD	Definition instruction	Near	Researcher-generated comprehension test	1.71***	N/A
	Definition instruction	Near	Researcher-generated vocabulary test	1.49***	N/A
Bos, Anders, Filip, & Jaffe (1989); 50 high school students with LD	Dictionary method	Near	Researcher-generated comprehension test	1.67***	.00
SM, SFA, SSFA Combination					
Bos & Anders (1992); Study 1: 42 upper-elementary students with LD	Definition instruction	Near	Researcher-generated comprehension test	.81	.86
	Definition instruction	Near	Researcher-generated vocabulary test	.50	1.15
Bos & Anders (1992); Study 3: 47 upper-elementary students with LD	Definition instruction	Near	Researcher-generated comprehension test	1.46	N/A
	Definition instruction	Near	Researcher-generated vocabulary test	1.28	N/A
Visual Display					
Darch & Carnine (1986); 24 upper-elementary students with LD	Text-only	Near	Researcher-generated comprehension test	1.73***	N/A
	Text-only	Far	Researcher-generated far-transfer comprehension Test	.64	N/A

continued next page

Table 1 continued

Independent Variables, Dependent Measures, and Effect Sizes of Individual Experiments

Subject/GO Type/Study/ Participants	Control Condition	Near or Far Transfer	Dependent Measure	Effect Size Posttest *M*	Maintenance *M*
SOCIAL STUDIES continued					
Visual Display continued					
Darch, Carnine, & Kame'enui (1986); 84 junior high school students with LD	Directed reading	Near	Researcher-generated comprehension test	.79***	N/A
	Directed reading	Far	Researcher-generated far-transfer comprehension test	.82***	N/A
DiCecco & Gleason (2002); 24 junior high students with LD	Guided discussion	Near	Researcher-generated fact quizzes	.05	N/A
	Guided discussion	Far	Relational statements in two written essays	.94***	N/A
Hudson (1996); 21 junior high students with LD	Note-taking guide	Near	Researcher-generated knowledge test	1.58***	2.37***
	Note-taking guide	Far	Researcher-generated inference test	1.00**	2.08***
SCIENCE					
Semantic Mapping					
Bos & Anders (1990); 61 junior high students with LD	Definition instruction	Near	Researcher-generated vocabulary test	1.27***	.94**
	Definition instruction	Near	Researcher-generated comprehension test	1.33***	.54
Bos & Anders (1992); Study 6: 22 junior high students with LD	Normative	Near	Researcher-generated comprehension test	2.16***	.78*
Reyes, Gallego, Duran, & Scanlon (1989); 61 junior high students with LD	Definition instruction	Near	Researcher-generated comprehension test	1.27***	.94**
	Definition instruction	Near	Researcher-generated vocabulary test	1.33***	.47

continued next page

Table 1 continued

Independent Variables, Dependent Measures, and Effect Sizes of Individual Experiments

Subject/GO Type/Study/ Participants	Control Condition	Near or Far Transfer	Dependent Measure	Effect Size Posttest *M*	Maintenance *M*
SCIENCE continued					
Semantic Feature Analysis					
Bos & Anders (1990 61 junior high students with LD	Definition); instruction	Near	Researcher-generated vocabulary test	1.03**	.66*
	Definition instruction	Near	Researcher-generated comprehension test	1.47***	.44
Bos & Anders (1992); Study 6: 22 junior high students with LD	Normative	Near	Researcher-generated comprehension test	.17	.73*
Reyes, Gallego, Duran, & Scanlon (1989); 61 junior high students with LD	Definition instruction	Near	Researcher-generated comprehension test	1.03***	.66*
	Definition instruction	Near	Researcher-generated vocabulary test	1.47***	.44
Syntactic/Semantic Feature Analysis					
Bos & Anders (1990); 61 junior high students with LD	Definition instruction	Near	Researcher-generated vocabulary test	.64	1.38***
	Definition instruction	Near	Researcher-generated comprehension test	1.18**	1.40***
Reyes, Gallego, Duran, & Scanlon (1989); 61 junior high students with LD	Definition instruction	Near	Researcher-generated comprehension test	.64	1.38***
	Definition instruction	Near	Researcher-generated vocabulary test	1.18***	1.40***
SM/SFA/SSFA Combination					
Bos & Anders (1992); Study 2: 61 junior high students with LD	Definition instruction	Near	Researcher-generated comprehension test	1.22	.78
	Definition instruction	Near	Researcher-generated vocabulary test	.92	1.01
Bos & Anders (1992); Study 4: 53 junior high students with LD	Definition instruction	Near	Researcher-generated comprehension test	1.51	1.51
	Definition instruction	Near	Researcher-generated vocabulary test	.83	.79

continued next page

Table 1 continued

Independent Variables, Dependent Measures, and Effect Sizes of Individual Experiments

Subject/GO Type/Study/ Participants	Control Condition	Near or Far Transfer	Dependent Measure	Effect Size Posttest *M*	Maintenance *M*
SCIENCE continued					
Visual Display					
Darch & Eaves (1986); 22 high school students with LD	Text-only	Near	Researcher-generated comprehension test	1.29***	.35
	Text-only	Near	Researcher-generated far-transfer comprehension test	.64	N/A
Griffin, Simmons, & Kame'enui (1991); 28 upper-elementary students with LD	No visual display	Near	Researcher-generated comprehension test	.51	N/A
SCIENCE/SOCIAL STUDIES					
Visual Display					
Darch & Gersten (1986); 24 high school students with LD	Basal instruction	Near	Researcher-generated comprehension test	1.72***	N/A
Semantic Feature Analysis					
Darch & Gersten (1986); 24 high school students with LD	Dictionary method	Near	Researcher-generated vocabulary test	1.66***	.00
ENGLISH/LANGUAGE ARTS					
Cognitive Mapping					
Boyle (1996); 30 junior high school students with LD or MMR	Typical reading instruction	Near	Formal reading inventory	.34	N/A
	Typical reading instruction	Near	Below-grade-level literal comprehension test	.87**	N/A
	Typical reading instruction	Near	On-grade-level literal comprehension test	1.33***	N/A
	Typical reading instruction	Far	Below-grade-level inferential comprehension test	.76**	N/A
	Typical reading instruction	Far	On-grade-level inferential comprehension test	.96***	N/A

continued next page

Table 1 continued

Independent Variables, Dependent Measures, and Effect Sizes of Individual Experiments

Subject/GO Type/Study/ Participants	Control Condition	Near or Far Transfer	Dependent Measure	Effect Size Posttest *M*	Maintenance *M*
ENGLISH/LANGUAGE ARTS continued					
Cognitive Mapping					
Mastropieri & Peters (2003); 20 junior high students with LD	List illustration	Near	Research-generated featured item recall test	1.42***	N/A
	List illustration	Far	Research-generated featured item recall test	1.06**	N/A
Semantic Mapping					
Englert & Mariage (1991); 28 upper-elementary students with LD	Typical reading instruction	Near	Written free recall	1.84***	N/A
Visual Display					
Boyle (2000); 24 high school students with LD or MMR	No training	Near	Literal comprehension test	1.14***	N/A
	No training	Near	Relational comprehension test	.91**	N/A
	No training	Far	Inferential comprehension test	.51	N/A
MATHEMATICS					
Visual Display					
Ives (2007); Study 1: 30 high school students with LD	Control	Near	Teacher-generated test	.67*	N/A
	Control	Near	Researcher-generated test (concepts)	1.06**	.94**
	Control	Far	Researcher-generated test (system solving)	.16	.00
Ives (2007); Study 2: 20 high school students with LD	Control	Near	Researcher-generated test	.49	N/A

***$p < .001$; **$p < .05$; *$p < .1$; all GOs were created by the researchers.

Generally, instruction for the experimental groups included one to two sessions focused solely on how to use the GO, one to two sessions of prompted practice using the GO, and independent student use of the GO for the remainder of sessions. During the initial sessions, the teacher or researcher presented the GO to students and described how it illustrated relationships. For example, Darch and Carnine (1986) presented their visual display via an overhead projector, and students followed along while the teacher used a script to describe the various cells in the display and the interrelationships between them.

The following sessions generally included the instructor explicitly guiding the students in creating or filling out the GO. For example, Bos and Anders (1990) explicitly prompted the students in each step of creating a hierarchical semantic map from a vocabulary list. This level of assistance was gradually faded. For instance, Darch and Gersten (1986) first presented a visual display with all the cells labeled, and prompted the entire group in answering questions about specific facts in the GO. The researchers followed this by guiding the students through a visual display that did not provide cell labels. Finally, individual students were prompted in labeling blank visual displays. Instruction in the remaining sessions generally focused on independent use of the GO by the students, in addition to text or lecture presentations. However, in each of the visual display studies all of the content was presented solely through the GO.

Each of the interventions lasted between one and seven weeks, with an additional one to four weeks between posttest and maintenance measures. All of the studies were conducted in a resource classroom during or after the regular school day.

What Are the Overall Effects of GOs on the Posttest Performance of Students With LD?

After the removal of six outliers, there was a large overall standardized effect of GOs on the posttest performance (e.g., multiple-choice comprehension, vocabulary, written recall) of students with LD across all studies ($\overline{ES} = .91$, $SE = .062$) for both random- and fixed-effects models and a 95% confidence interval of .79, 1.03 for the random effects model. Table 2 provides the full comparison between the random- and fixed-effects models.

Do These Effects Maintain Over Time?

Twenty-nine studies included maintenance measures. In each of the studies, measures consisted of multiple-choice comprehension or vocabulary items. These measures were given to students one to four weeks after the conclusion of the intervention.

The test of homogeneity for overall maintenance effects produced a nonsignificant Q statistic ($Q = 24.49$, $cv = 35.17$). Therefore, similar to the overall posttest effects, results of both the random- and fixed-effects models are reported. After removal of five outliers, there was a medium overall effect for maintenance across all studies ($\overline{ES} = .56$, $SE = .074$) with a 95% confidence interval of .41, .70 for the random-effects model. Table 3 provides the full comparison between the random- and fixed-effects models.

Differential Effects

Based on number of effect sizes and significant Q-values, the remaining analyses of differential effects are

Table 2
Overall $\overline{ES}$ for Fixed and Random Models

	$\overline{ES}$	*SE* of $\overline{ES}$	*Z*-test	95% CI Lower	95% CI Upper
Fixed Model	.9127	.062	14.677*	.79	1.03
Random Model	.9061	.062	14.615*	.78	1.02

*$p < .001$.

Table 3
Maintenance $\overline{ES}$ for Fixed and Random Models

				95% CI	
	$\overline{ES}$	*SE of* $\overline{ES}$	*Z*-test	Lower	Upper
Fixed Model	.5625	.077	7.305*	.41	.71
Random Model	.5614	.074	7.586*	.41	.70

$^*p < .001$.

reported as the random-effects model. This provides a more conservative estimate of effect.

Are There Differential Effects by Type of Measure (Near or Far Transfer)? Forty-five individual estimates of effect were calculated for near-transfer measures at posttest and 27 individual estimates of effect for maintenance. In all but two articles, near-transfer measures consisted of researcher-generated multiple-choice questions on material directly covered in the lessons. In the remaining two articles, Boyle (1996) used a standardized measure of reading comprehension and Englert and Mariage (1991) used a measure of written free recall, respectively. Overall mean effect size was 1.07 for posttest and .78 for maintenance.

Ten individual estimates of effect were calculated for far-transfer measures at posttest and two individual estimates of effect for maintenance. These measures tested students' ability to apply knowledge to situations not directly covered in the text or lecture. For example, Hudson (1996) included the question "Name one way the environment influenced the culture of the Arctic tribes" (p. 81). The teacher never stated the causal relation between environment and culture; only facts about environment and culture were stated.

Table 4
Near- and Far-Transfer – Random-Effects Model

	Posttest			Maintenance		
		95% CI			95% CI	
	$\overline{ES}$	Lower	Upper	$\overline{ES}$	Lower	Upper
Near Transfer	1.065	.94 *n* = 45	1.19	.7809	.63 *n* = 27	.93
Far Transfer	.6127	.36 *n* = 10	.87	.6886	.07 *n* = 2	1.31

Note. n = number of *ESs*.

Table 5
Type of Graphic Organizer – Random-Effects Model

	Posttest			Maintenance		
		95% CI			95% CI	
	$\overline{ES}$	Lower	Upper	$\overline{ES}$	Lower	Upper
Cognitive Mapping	.8914	.58 *n* = 7	1.21	N/A	N/A *n* = 0	N/A
Semantic Mapping	1.251	.94 *n* = 7	1.56	.6925	.38 *n* = 6	1.01
Semantic Feature Analysis	1.187	1.06 *n* = 10	1.32	.3695	.23 *n* = 8	.51
Syntactic/Semantic Feature Analysis	.91	.82 *n* = 4	.99	1.39	.95 *n* = 4	1.83
SM/SFA/SSFA Combination	1.062	.93 *n* = 8	1.19	1.013	.86 *n* = 6	1.17
Visual Display	.7486	.56 *n* = 19	.93	.7877	.42 *n* = 5	1.16

Note. n = number of *ESs*.

As for near-transfer, far-transfer measures consisted of researcher-generated multiple-choice questions in all but one study. Boyle (1996) used a standardized measure for far transfer. The overall mean effect size was .61 for posttest and .69 for maintenance. Table 4 provides the full comparison between near- and far-transfer measures.

Are There Differential Effects by Type of GO? The types of GOs used in the studies matched the definitions in the introduction to this analysis (e.g., cognitive mapping, SM, SFA, SSFA, visual display). However, in one article containing eight studies (Bos & Anders, 1992), the researchers used a combination of SM, SFA, and SSFA. The method they utilized to present their results prohibited disaggregation of the findings. Therefore, a sixth category (SM/SFA/SSFA Combination) was added to the analysis. Large posttest effects (e.g., .74-1.2) were found for all types of GOs except visual displays. Visual displays had a moderate effect (e.g., .74). There were no statistically significant differences between GOs with large posttest effects. For maintenance measures, SSFA and SM/SFA/SSFA Combination had significantly larger effects than the other GO types (e.g., 1.39, 1.01). Table 5 provides the full comparison between types of GOs.

Are There Differential Effects by Subject Area? Posttest effects were calculated for the subject areas of English/writing/reading, mathematics, science, and social studies. Large posttest effects were found for all subject (e.g., .96-1.05) areas except mathematics (e.g., .59). Mathematics had a moderate posttest effect that was significantly smaller than the other subject areas. Maintenance effects were calculated for mathematics, science, and social studies. Science had a large maintenance effect (e.g., .80) that was significantly larger than the moderate effects for mathematics and social studies. Table 6 provides the full comparison between GOs by subject area.

Are There Differential Effects by Stage of Attending to Verbal Material (Before, During, After Instruction)? There was not enough information to quantify differential effects by stage of instruction. All but one study included GOs before and during instruction, and not enough information was provided to disaggregate these data. One study, Englert and Mariage (1991), used GOs after instruction. The unstandardized effect size for a near-transfer, written free recall measure was large (*ES* = 1.84).

Discussion

As was the casae in previous research syntheses (e.g., Gajria et al., 2007; Kim et al., 2004; Moore & Readence, 1984), findings from this meta-analysis indicate that GOs improve the factual comprehension of upper-elementary, intermediate, and secondary students with LD. Unlike these previous reviews, this analysis also indicates that GOs may improve vocabulary and inference/relational comprehension for students with LD.

What Are the Overall Effects of GOs on the Posttest Performance of Students With LD?

Overall, there was a large mean effect for posttest performance (ES = .91, SE = .06) of students with LD using both the fixed-effects and random-effects models. The immediate posttest performance spanned multiple constructs, including multiple-choice factual comprehension, vocabulary, and written recall requiring relational comprehension. This suggests that GOs are effective in not only improving basic skills (e.g., factual recall) but also higher-level skills (e.g., inference). This finding is consistent with the theories of Ausubel (1968) and Mayer (1979) that GOs may especially assist lower ability learners in both basic and higher level skills by creating an easier context for assimilating information into their memory.

Do These Effects Maintain Over Time?

There was a moderate mean effect for maintenance (*ES* = .56, *SE* = .07) of students with LD. The significant drop-off from posttest to maintenance is consistent with the findings of the other GO research syntheses. The drop has been attributed to lack of clarity regarding the duration and length of intervention sessions needed to positively affect maintenance (Gajria et al., 2007; Gersten, Fuchs, Williams, & Baker, 2001). The relatively short duration of the intervention studies (e.g., 1-7 weeks) may not have provided sufficient instruction time for students to use GOs independently. However, a closer look at effects by type of GO shows that effect sizes for visual displays and SSFA were larger for maintenance than posttest. This may lend support to the visual argument hypothesis (Waller, 1981), which posits that GOs that are structured in a way that facilitates understanding and perception of concept relationships are superior to more complicated GOs that may require instruction to recognize conceptual relationships (Dexter, 2010). The visual displays and SSFA were more computationally efficient than the other GOs. That is,

Table 6
Subject Area – Random-Effects Model

	Posttest			Maintenance		
		95% CI			95% CI	
	$\overline{ES}$	Lower	Upper	$\overline{ES}$	Lower	Upper
English/Writing/	.9612	.72 *n* = 11	1.20	N/A	N/A *n* = 0	N/A
Mathematics	.5942	.21 *n* = 4	.98	.4559	.07 *n* = 2	.99
Science	1.052	.88 *n* = 23	1.23	.8035	.64 *n* = 20	.97
Social Studies	1.037	.85 *n* = 19	1.22	.6535	.38 *n* = 8	1.03

Note. n = number of *ESs*.

they were simple enough for students to recognize conceptual relationships without teacher instruction. This may explain why maintenance effects were larger for these types of GOs.

Are There Differential Effects by Type of Measure (Near or Far Transfer)?

This meta-analysis also separated results into near-transfer and far-transfer measures. Near-transfer results (i.e., measures applying knowledge to situations directly covered in the text or lecture) indicate that GOs are effective strategies for improving factual recall, factual and relational comprehension, and vocabulary knowledge. Across all near-transfer studies, the mean effect size was large (*ES* = 1.07), and maintenance effects were moderate (*ES* = .78). Students using GOs significantly outperformed their peers receiving typical classroom instruction on near-transfer measures. Interestingly, more complicated GOs requiring intensive teacher instruction (e.g., SM, SFA) resulted in the largest effects for near-transfer posttest measures. This indicates that while these GOs are difficult to understand independently, with appropriate instruction they are superior to less complicated GOs for immediate factual recall.

Far-transfer results (i.e., measures applying knowledge to situations not directly covered in the text or lecture) indicate that GOs may also improve the inference skills and relational knowledge of secondary students with LD. Across all far-transfer studies, the mean effect size was moderate (*ES* = .61), and maintenance effects were moderate (*ES* = .69).

It is interesting to note that for far-transfer measures, maintenance effect sizes were larger than posttest effect sizes. Previous research has indicated students with LD typically perform poorly on far-transfer tasks due to their inability to detect underlying concepts in verbal information due to difficulty assimilating verbal information with previous knowledge (Suritsky & Hughes, 1991). The results of this analysis demonstrate that GOs may bridge the gap between verbal information and prior knowledge and assist students with LD in far-transfer tasks. This finding supports Mayer's (1979) assimilation theory, which posits that GOs that assimilate material to a broader set of past experiences allow superior transfer to new situations.

The finding is also consistent with the research of Robinson and colleagues (Robinson et al., 1998; Robinson & Kiewra, 1995; Robinson & Schraw, 1994; Robinson & Skinner, 1996), comparing visual displays (e.g., tree diagrams, matrices, network charts) with traditional, non-graphic outlines. In each of the studies, groups of nondisabled college students using GOs and traditional outlines equally outperformed text-only groups in factual recall, but the GO groups significantly outperformed the outline and text-only groups in identifying concept relations and making far-transfer concept comparisons.

Are There Differential Effects by Subject Area?

This analysis also examined the effects of GOs by subject area. All subject areas had moderate to large effects for posttest and maintenance measures. The largest effects were in science (*ES* = 1.05 for posttest, *ES* = .80 for maintenance) and the smallest effects in mathematics (*ES* = .59 for posttest, *ES* = .46 for maintenance). The large effects in science may be explained by the unfamiliar, technical vocabulary and content often based on relationships between concepts (Lovitt et al., 1986). This type of content lends itself to computationally efficient GOs that make relationships explicit and clear. Also, students may rely more heavily on content enhancements like GOs when content seems strange or foreign.

The small effects for mathematics may be explained by the fact that the information was much more abstract than the other subject areas. It may also be explained by the fact that the mathematics study used a visual display, which only had a moderate overall effect. Use of GOs with mathematics concepts and solving systems of linear equations are only beginning in the field. Ives (2007) offerred several implications for future research based on his initial study in this field. It will take time and more study to fully understand the effects of GOs on mathematics understanding.

Are There Differential Effects by Stage of Attending to Verbal Material (Before, During, After Instruction)?

Finally, a previous GO research synthesis (Moore & Readence, 1984) reported that GOs presented as text summaries after instruction were more effective than GOs presented before or during instruction. This finding cannot be corroborated by the present analysis because only one study (Englert & Mariage, 1991) used GOs after instruction. While this study had an extremely large effect size (1.84), more studies are needed to confirm this finding. The current analysis points to effective instruction and choice of GO to be more important than stage of attending to verbal material in effectiveness of the intervention.

Methodological Limitations

There are two methodological limitations to the conclusions of this analysis. First, there is the possibility of a publication bias. According to Smith (1980), as well as Lipsey and Wilson (1993), published articles have a larger mean effect size than unpublished studies. This bias, also known as a "file-drawer" effect, happens because studies with null findings are less likely to be published by major journals, often leaving offending data in a researcher's file drawer (Lipsey & Wilson, 2001). In this analysis, based on the advice of Slavin (1995), we purposefully selected only published studies to ensure the highest quality of research designs. While this eliminated our ability to compare mean effect sizes of published studies versus unpublished studies, we were able to utilize Rosenthal's (1979) fail-safe N statistic, which was later adapted by Orwin (1983). This statistic determines "the number of studies with an effect size of zero needed to reduce the mean effect size to a specified or criterion level" (Lipsey & Wilson, 2001, p. 166).

The formula is as follows:

$$k_0 = k\left[\frac{\overline{ES}_k}{\overline{ES}_c} - 1\right],$$

where k_0 is the number of effect sizes of zero needed to reduce the mean effect size to $\overline{ES}_c$ (criterion effect level), k is the number of current studies, and $\overline{ES}_k$ is the current weighted mean effect size. Using this statistic, reduction of our overall weighted effect size of .91 to .50 would take 45 additional unpublished studies with an effect size of zero. While this may be a possibility, it is unlikely that 45 additional null studies exist in researchers' file drawers.

Second, our 55 unique effect sizes or studies were culled from only 16 published articles. While this is an acceptable practice (Lipsey & Wilson, 2001), it does limit the generalizability of the findings because there were only 21 distinct samples of students with LD (total $N = 808$). Therefore, caution should be exercised in generalizing these findings to all intermediate and secondary students with LD.

Individual Study Limitations

Three limitations to the individual studies warrant consideration. First, while all of the effect sizes in the present analysis were based on differences between a treatment group and a control group, it is not clear if the control conditions constituted an adequate standard to measure the effects of GO interventions (Gersten, Baker, & Lloyd, 2000). The control conditions in the studies used primarily typical classroom practices (e.g., dictionary instruction) rather than more closely comparable practices (e.g., outlines, structured overviews). While this provides much evidence for GO effects over typical classroom practice, it does not yield information comparing GOs with other researched practices (Kim et al., 2004)

Second, while the results indicate large effects for vocabulary, inference, and comprehension, it is important to note that all but one study used measures that were researcher-created and closely tied to the content. While these measures should have good content validity, there is no way to measure broader construct validity. Only Boyle (1996) used a standardized measure for reading comprehension. This fact may limit the generalizability of these findings and questions the actual level of understanding obtained by students in the GO conditions.

Finally, studies using Content Enhancement Routines were not included in this meta-analysis because their effects could not be attributed solely to the GO, even though GOs are a vital component of each routine. Research on Content Enhancement Routines has been conducted for over two decades, focusing primarily on assisting all students, including those with LD, thrive in the often rigorous intermediate and secondary general education content classrooms (Bulgren, 2006; Bulgren, Deshler, & Lenz, 2007 These routines have been shown to improve both basic skills (e.g., factual knowledge, comprehension) and higher-order skills (e.g., manipulation, extension, generalization) for students with LD (e.g., Bulgren et al., 2000; Bulgren et al., 2002; Bulgren et al., 2009; Scanlon et al., 1996). While results of these studies did not fit into our analysis, they do serve as examples of effective interventions utilizing GOs as a component of a larger routine.

Implications for Practice

The major implication of this study for applied practice is consistent with assimilation theory and the visual argument hypothesis; that is, more instruction-intensive types of GOs (e.g., SM, SFA) are better for immediate factual recall while more computationally efficient GOs (e.g., visual display, SSFA) are better for maintenance and transfer. This knowledge can help teachers in designing GOs for initial instruction and for re-teaching, studying, and retention purposes. For instance, a semantic map for initial instruction, followed by a simpler visual display for review and study will potentially maximize the effects of recall, maintenance, and far-transfer for students with LD.

Another implication for practice is that, regardless of GO type, a teacher must explicitly teach students how to use a given GO. Students with LD need explicit instruction to understand how concepts are related, to recognize differences between main and subordinate ideas, and to put all the pieces together to make a clear picture of the content being learned no matter how implicit a GO may seem. A teacher's use of effective instruction practices (e.g., modeling, corrective feedback, etc.) will positively impact the intervention's effectiveness.

Conclusions and Implications for Future Research

This meta-analysis found that, in comparison to activities such as reading text passages, attending to lectures, and participating in typical classroom practice (e.g., dictionary instruction), GOs are more effective on posttest, maintenance, and transfer measures. However, this conclusion must be tempered for several reasons.

First, each of the studies took place in self-contained resource classrooms. This may not be typical for today's upper-elementary, intermediate, and secondary students with LD. As many students with LD are now fully included with nondisabled peers in core content classes, it is important to closely examine how GOs will work in this setting. The feasibility and practicality of GOs must be examined in general education settings and recommendations for effective use put forth.

Second, there is great need for GO replication studies. Only three articles in the present meta-analysis were published in the past 10 years. More current group design, randomized control trials, is needed to fully validate the benefits of GOs across all secondary students with LD.

Finally, for student independent practice, it was not always clear from the studies if the GO was used correctly or at all. For instance, when students independently filled in a blank GO, there was no reported procedure for ascertaining if they were properly labeling main and subordinate details. Likewise, several of the studies (e.g., Anders, Bos, & Filip, 1984; Bos & Anders, 1990; Bos & Anders, 1992) reported students had a GO and text to study for the posttests. They did not include a procedure for making sure the students were actually using the GO to study. These students may have been using the text as their study guide. This lack of control may somewhat negate the attribution of effects to the GO. Future research must tightly control for these potential problem areas.

Taking the above issues into account, the evidence in this analysis should still persuade educational practitioners to make well-planned and well-instructed use of GOs. There were no significant negative effects across any of the categories of analysis and no other identified detrimental effect. A thoughtful combination of types of GOs will help make the learning process more efficient for upper-elementary, intermediate, and secondary students with LD.

References

*Anders, P. L., Bos, C. S., & Filip, D. (1984). The effect of semantic feature analysis on the reading comprehension of learning disabled students. In J. A. Niles & L. A. Harris (Eds.), *Changing perspectives on reading/language processing and instruction* (pp. 162-166). Rochester, NY: National Reading Conference.

Armbruster, B. B. (1984). The problems of "inconsiderate" text. In G. G. Duffy, L. R. Roehler, & J. M. Mason (Eds.), *Comprehension instruction: Perspectives and suggestions* (pp. 202-217). New York, NY: Longman.

Ausubel, D. P. (1968). *Educational psychology: A cognitive view.* New York, NY: Holt, Rinehart, & Winston.

Baumann, J. F. (1984). The effectiveness of a direct instruction paradigm for teaching main idea comprehension. *Reading Research Quarterly, 20,* 93-115.

Bergerud, D., Lovitt, T. C., & Horton, S. (1988). The effectiveness of textbook adaptations in life science for high school students with learning disabilities. *Journal of Learning Disabilities, 21,* 70-76.

Boon, R. T., Fore III, C., Ayres, K., & Spencer, V. G. (2005). The effects of cognitive organizers to facilitate content-area learning for students with mild disabilities: A pilot study. *Journal of Instructional Psychology, 32,* 101-117.

*Bos, C. S., & Anders, P. L. (1990). Effects of interactive vocabulary instruction on the vocabulary learning and reading compre-hension of junior-high learning disabled students. *Learning Disability Quarterly, 13,* 31-42.

*Bos, C. S., & Anders, P. L. (1992). Using interactive teaching and learning strategies to promote text comprehension and content learning for students with learning disabilities. *International Journal of Disability, Development, and Education, 39,* 225-238.

*Bos, C. S., Anders, P. L., Filip, D., & Jaffe, L. E. (1989). The effects of an interactive instructional strategy for enhancing reading comprehension and content area learning for students with learning disabilities. *Journal of Learning Disabilities, 22,* 384-389.

Bos, C. S., & Vaughn, S. (2002). *Strategies for teaching students with learning and behavior problems* (5th ed.). Boston, MA: Allyn & Bacon.

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

*Boyle, J. R. (2000). The effects of a Venn diagram strategy on the literal, inferential, and relational comprehension of students with mild disabilities. *Learning Disabilities: A Multidisciplinary Journal, 10,* 5-13.

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

Boyle, J. R., & Yeager, N. (1997). Blueprints for learning: Using cognitive frameworks for understanding. *Teaching Exceptional Children, 29,* 26-31.

Brigham, F. J., Scruggs, T. E., & Mastropieri, M. A. (1995). Elaborative maps for enhanced learning of historical information: Uniting spatial, verbal, and imaginable information. *The Journal of Special Education, 28,* 440-460.

Bryant, D. P., Ugel, N., Thompson, S., & Hamff, A. (1999). Instructional strategies for content-area reading instruction. *Intervention in School and Clinic, 34,* 293-302.

Bulgren, J. A. (2006). Integrated content enhancement routines: Responding to the needs of adolescents with disabilities in rigorous inclusive secondary content classes. *Teaching Exceptional Children, 38,* 54-58.

Bulgren, J. A., Deshler, D. D., Schumaker, J. B., & Lenz, B. K. (2000). The use and effectiveness of analogical instruction in diverse secondary content classrooms. *Journal of Educational Psychology, 92,* 426-441.

Bulgren, J., Deshler, D. D., & Lenz, B. K. (2007). Engaging adolescents with LD in higher order thinking about history concepts using integrated content enhancement routines. *Journal of Learning Disabilities, 40,* 121-133.

Bulgren, J. A., Lenz, B. K., Schumaker, J. B., Deshler, D. D., & Marquis, J. G. (2002). The use and effectiveness of a comparison routine in diverse secondary content classrooms. *Journal of Educational Psychology, 94,* 356-371.

Bulgren, J. A., Marquis, J. G., Lenz, B. K., Schumaker, J. B., & Deshler, D. D. (2009). Effectiveness of question exploration to enhance students' written expression of content knowledge and comprehension. *Reading and Writing Quarterly, 25,* 271-289.

Bulgren, J. A., Schumaker, J. B., & Deshler, D. D. (1988). Effectiveness of a concept teaching routine in enhancing the performance of LD students in secondary-level mainstream classes. *Learning Disabilities Quarterly, 11,* 3-17.

Burns, M. (1994). Empirical analysis of drill ratio research. *Remedial and Special Education, 25,* 167-173.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

*Darch, C., & Carnine, D. (1986). Teaching content area material to learning disabled students. *Exceptional Children, 53,* 240-246.

*Darch, C. B., Carnine, D. W., & Kame'enui, E. J. (1986). The role of graphic organizers and social structure in content area instruction. *Journal of Reading Behavior, 18,* 275-295.

*Darch, C., & Eaves, R. (1986). Visual displays to increase comprehension of high school learning-disabled students. *The Journal of Special Education, 20,* 309-318.

*Darch, C., & Gersten, R. (1986). Direction-setting activities in reading comprehension: A comparison of two approaches. *Learning Disability Quarterly, 9,* 235-243.

Deshler, D. D., Ellis, E., & Lenz, B. K. (1996). *Teaching adolescents with learning disabilities: Strategies and methods.* Denver, CO: Love.

Dexter, D. D. (2010). Graphic organizers and their effectiveness for students with learning disabilities. *Thalamus, 26,* 51-67.

*DiCecco, V. M., & Gleason, M. M. (2002). Using graphic organizers to attain relational knowledge from expository text. *Journal of Learning Disabilities, 35,* 306-320.

*Englert, C. S., & Mariage, T. V. (1991). Making students partners in the comprehension process: Organizing the reading POSSE. *Learning Disability Quarterly, 14,* 123-138.

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2007). *Learning disabilities: From identification to intervention.* New York, NY: The Guilford Press.

Gajria, M., Jitendra, A. K., Sood, S., & Sacks, G. (2007). Improving comprehension of expository text in students with LD: A research synthesis. *Journal of Learning Disabilities, 40,* 210-225.

Gardill, M. C., & Jitendra, A. K. (1999). Advanced story map instruction: Effects on the reading comprehension of students with learning disabilities. *The Journal of Special Education, 33,* 2-17.

Gersten, R., Baker, S., & Lloyd, J. W. (2000). Designing high-quality research in special education: Group experimental design. *The Journal of Special Education, 34,* 2-18.

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

*Griffin, C. C., Simmons, D. C., & Kame'enui, E. J. (1991). Investigating the effectiveness of graphic organizers instruction on the comprehension and recall of science content by students with learning disabilities. *Reading, Writing, and Learning Disabilities, 7,* 355-376.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107-128.

Hedges, L. V., Shymansky, J. A., & Woodworth, G. (1989). *A practical guide to modern methods of meta-analysis.* Washington, DC: National Science Teachers Association.

Holmes, B. C. (1985). The effects of a strategy and sequenced material on the inferential comprehension of disabled readers. *Journal of Learning Disabilities, 18,* 542-546.

Horton, P. B., McConney, A. A., Gallo, M., Woods, A. L., Senn, G. J., & Hamelin, D. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education, 77,* 95-111.

Horton, S. V., Lovitt, T. C., & Bergerud, D. (1990). The effectiveness of graphic organizers for three classifications of secondary students in content area classes. *Journal of Learning Disabilities, 23,* 12-22.

*Hudson, P. (1996). Using a learning set to increase the test performance of students with learning disabilities in social studies classes. *Learning Disabilities Research & Practice, 11,* 78-85.

Hughes, C. A., Maccini, P., & Gagnon, J. C. (2003). Interventions that positively impact the performance of students with learning disabilities in secondary general education classes. *Learning Disabilities, 12,* 101-111.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings.* Thousand Oaks, CA: Sage Publications, Inc.

Idol, L., & Croll, V. J. (1987). Story-mapping training as a means of improving reading comprehension. *Learning Disability Quarterly, 10,* 214-229.

*Ives, B. (2007). Graphic organizers applied to secondary algebra instruction for students with learning disorders. *Learning Disabilities Research & Practice, 22,* 110-118.

Ives, B., & Hoy, C. (2003). Graphic organizers applied to higher-level secondary mathematics. *Learning Disabilities Research & Practice, 18,* 36-51.

Johnson, L., Graham, S., & Harris, K. R. (1997). The effects of goal setting and self-instructions on learning a reading comprehension strategy: A study with students with learning disabilities. *Journal of Learning Disabilities, 30,* 80-91.

Kim, A., Vaughn, S., Wanzek, J., & Wei, S. (2004). Graphic organizers and their effects on the reading comprehension of students with LD: A synthesis of research. *Journal of Learning Disabilities, 37,* 105-118.

Kulhavy, R. W., Stock, W. A., & Caterino, L. C. (1994). Reference maps as a framework for remembering text. In W. Schnotz & R. W. Kulhavy (Eds.). *Comprehension of graphics* (pp. 153-162). New York, NY: Elsevier Science.

Lenz, B. K., Adams, G. L., Bulgren, J. A., Pouliot, N., & Laraux, M. (2007). Effects of curriculum maps and guiding questions on the test performance of adolescents with learning disabilities. *Learning Disability Quarterly, 30,* 235-244.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment. *The American Psychologist, 48,* 1181-1201.

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

Lovitt, T., Rudsit, J., Jenkins, J., Pious, C., & Benedetti, D. (1986). Adapting science materials for regular and learning disabled seventh graders. *Remedial and Special Education, 7,* 31-39.

Mastropieri, M. A., Scruggs, T. E., & Levin, J. R. (1987). Learning-disabled students' memory for expository prose: Mnemonic versus nonmnemonic pictures. *American Educational Research Journal, 24,* 505-519.

Mayer, R. E. (1979). Can advanced organizers influence meaningful learning? *Review of Educational Research, 49,* 371-383.

Minskoff, E., & Allsopp, D. (2003). *Academic success strategies for adolescents with learning disabilities and ADHD.* Baltimore, MD: Brookes Publishing.

Moore, D. W., & Readence, J. F. (1984). A quantitative and qualitative review of graphic organizer research. *Journal of Educational Research, 78,* 11-17.

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76,* 413- 448.

Orwin, R. G. (1983). A fail-safe *N* for effect size in meta-analysis. *Journal of Educational Statistics, 8,* 157-159.

Pearson, P. D., & Johnson, D. (1978). *Teaching reading comprehension.* New York, NY: Holt, Rinehart, & Winston.

Pehrsson, R., & Denner, P. (1988). Semantic organizers: Implications for reading and writing. *Topics in Language Disorders, 8,* 24-37.

*Reyes, E. I., Gallego, M. A., Duran, G. Z., & Scanlon, D. J. (1989). Integration of internal concepts and external factors: Extending the knowledge of learning disabled adolescents. *The Journal of Early Adolescence, 9,* 112-124.

Rice, G. E. (1994). Need for explanations in graphic organizer research. *Reading Psychology: An International Quarterly, 15,* 39-67.

Rivera, D. P., & Smith, D. (1997). *Teaching students with learning and behavior problems* (3rd ed.). Boston, MD: Allyn & Bacon.

Robinson, D. H., Katayama, A. D., DuBois, N. F., & Devaney, T. (1998). Interactive effects of graphic organizers and delayed review on concept acquisition. *Journal of Experimental Education, 67,* 17-31.

Robinson, D. H., & Kierwa, K. A. (1995). Visual argument: Graphic organizers are superior to outlines in improving learning from text. *Journal of Educational Psychology, 87,* 455-476.

Robinson, D. H., & Schraw, G. (1994). Computational efficiency through visual argument: Do graphic organizers communicate relations in text too effectively? *Contemporary Educational Psychology, 19,* 399-415.

Robinson, D .H., & Skinner, C. H. (1996). Why graphic organizers facilitate search processes: Fewer words or computationally efficient indexing? *Contemporary Educational Psychology, 21,* 166-180.

Rosenthal, R. (1979). The "file-drawer" problem and tolerance for null results. *Psychological Bulletin, 86,* 638-641.

Rosenthal, R. (1994). Statistically describing and combining studies. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231-244). New York, NY: Russell Sage Foundation.

Scanlon, D., Deshler, D. D., & Schumaker, J. (1996). Can a strategy be taught and learned in secondary inclusive classrooms? *Learning Disabilities Research and Practice, 11,* 41-57.

Sinatra, R. C., Stahl-Gemake, J., & Berg, D. N. (1984). Improving reading comprehension of disabled readers through mapping. *The Reading Teacher, 38,* 22-29.

Slavin, R. E. (1995). Best evidence synthesis: An intelligent alternative to meta-analysis. *Journal of Clinical Epidemiology, 48,* 9-18.

Smith, M. L. (1980). Publication bias and meta-analysis. *Evaluation in Education, 4,* 22-24.

Sturm, J. M., & Rankin-Erickson, J. L. (2002). Effects of hand-drawn and computer-generated concept mapping on the expository writing of middle school students with learning disabilities. *Learning Disabilities Research & Practice, 17,* 124-139.

Suritsky, S. K., & Hughes, C. A. (1991). Benefits of notetaking: Implications for secondary and postsecondary students with learning disabilities. *Learning Disability Quarterly, 14,* 7-18.

Swanson, H. L., Hoskyn, M., & Lee, C. (1999). *Interventions for students with learning disabilities.* New York, NY: Guilford Press.

Thalheimer, W., & Cook, S. (2002, August). *How to calculate effect sizes from published research articles: A simplified methodology.* Retrieved from http://work-learning.com/effect_sizes.htm.

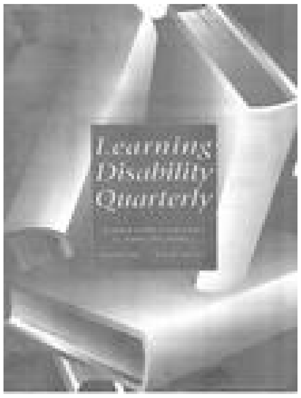
Waller, R. (1981, April). *Understanding network diagrams.* Paper presented at the annual meeting of the American Educational Research Association, Los Angeles.

Williams, J. P. (1993). Comprehension of students with or without learning disabilities: Identification of narrative themes and idiosyncratic text representations. *Journal of Educational Psychology, 85,* 631-641.

* = Studies included in meta-analysis.

Please address correspondence about this article to: Douglas D. Dexter, 230 CEDAR, The Pennsylvania State University, University Park, PA 16802; email:Ddd176@psu.edu

COPYRIGHT INFORMATION



Author: Dexter, Douglas D.; Hughes, Charles A.

Title: Graphic organizers and students with learning disabilities: a meta-analysis

Source: Learn Disabil Q 34 no1 Wint 2011 p. 51-72
ISSN: 0731-9487

Publisher: Council for Learning Disabilities
11184 Antioch, Suite 405, Overland Park, KS 66210