Dexter, D. (2012). Using Graphic Organizers to Teach Content Area Material to Students with Learning Disabilities. *International Journal for Research in Learning Disabilities, 1*(1), 79-105.

Using Graphic Organizers to Teach Content Area Material to Students with Learning Disabilities

Douglas D. Dexter[2]
The Pennsylvania State University

Abstract

A pretest-posttest comparison group design was used to investigate the effects of a semantic mapping lesson plus visual display versus a semantic mapping lesson alone on adolescents' with learning disabilities (LD) ability to gain and maintain factual knowledge from expository social studies material. In addition, a posttest only comparison group design was used to examine the effects of a semantic mapping lesson plus visual display versus a semantic mapping lesson alone on adolescents' with LD far-transfer ability. The results of this study supported the conclusion that semantic mapping was beneficial for factual recall, while the additive effect of a visual display significantly improved maintenance and far transfer for adolescents with LD. Results of this study also supported the conclusion that normally achieving students and low achieving students also benefit from semantic mapping and the visual display. This finding was consistent over written and multiple-choice measures. Implications for practice and future research are discussed.

The academic demands of the intermediate and secondary grades are escalated as material becomes more complex and abstract (Fletcher, Lyon, Fuchs, & Barnes, 2007). All students must use higher-order processing and comprehension skills to successfully navigate intermediate and secondary content curricula (Dexter & Hughes, 2011; Gajria, Jitendra, Sood, & Sacks, 2007; Graham & Hebert, 2010; Hughes, Maccini, & Gagnon, 2003), often through lecture and expository text presentation (Minskoff & Allsopp, 2003). The shift from primary to secondary grades is difficult for many students, but is especially so for 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 difficulties are exacerbated because of the complex structure of text and lectures at the secondary level, which are often

[2] Acknowledgement – I would like to sincerely thank Dr. Linda H. Mason for her support throughout this endeavor. This study would not have been possible without her expertise and guidance.

conceptually dense and filled with unfamiliar vocabulary (Gajria et al., 2007). 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 students in understanding increasingly abstract concepts (Bos & Vaughn, 2002; Dexter, 2010; Dexter & Hughes, 2011; Hughes et al., 2003; Ives & Hoy, 2003; Kim, Vaughn, Wanzek, &Wei, 2004; Nesbit & Adesope, 2007; Rivera & Smith, 1997).

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). GOs can be used before, during, and/or after a student attends to verbal (e.g., text or lecture) stimuli (Nesbit & Adesope, 2006).

Theoretical Framework for GOs with Students with LD

The theory of subsumption (Ausubel, 1960) and assimilation theory (Mayer, 1979) both offer direct implications about the possible benefits of GOs in learning. These two theories provide the basis for how GOs help facilitate understanding of unfamiliar material and clarify relationships between abstract concepts.

The research findings of both the theory of subsumption and assimilation theory appear to have specific implications for students with LD, although neither theory focused directly on this group of students (Dexter, 2010). Specifically, students with LD may benefit more from GOs than their non-disabled peers. A consistent pattern that emerged from the research on these theories is that students displaying lower verbal ability demonstrated larger gains than did students with average or high verbal ability, and these gains helped the students with lower verbal ability match the scores of peers with average verbal ability. Students with LD typically have low verbal ability (Kim et al., 2004) that often manifests itself as difficulty in connecting new material to prior knowledge (Williams, 1993). This is because, according to Mayer (1979), the specific structure of a GO may guide construction of cognitive structures in less knowledgeable students, but may conflict with pre-existing cognitive structures in more knowledgeable students.

Students with LD also typically perform poorly on far-transfer tasks (e.g., applying knowledge to new or unusual situations) due to their inability to detect underlying concepts in verbal information (Suritsky & Hughes, 1991). Based on the above theories, this may be due to difficulty assimilating verbal information with previous knowledge. The research evidence on assimilation theory suggests GOs may be the bridge in connecting verbal information with prior knowledge. This may dramatically assist students with LD in far-transfer tasks.

In addition, based on the visual argument hypothesis (Waller, 1981), Larkin and Simon (1987) concluded only "computationally efficient" (e.g., relationships more explicit than implicit) displays are effective for learning. Based on research published since Larkin and Simon's seminal work, other researchers have found patterns that support specific design principles that may achieve computational efficiency (McCrudden, Schraw, Lehman, & Poliquin, 2007; Robinson, Katayama, Dubois, & Devaney, 1998; Robinson & Kiewra, 1995; Robinson & Schraw, 1994; Robinson & Skinner, 1996). A general principle is that GOs are effective when they address the limitations of working memory in their design. This is consistent with the work of Swanson and Kim (2005), who found that students with LD performed significantly better on problem solving tasks when stress on working memory was minimized.

Research Base for GOs with Students with LD

Dexter and Hughes (2011) conducted a comprehensive meta-analysis of research studies examining GOs with secondary students with LD. Based on this meta-analysis, the major implication for applied practice is consistent with assimilation theory and the visual argument hypothesis: more instructionally intensive types of GOs (e.g., semantic maps) are better for immediate factual recall while more computationally efficient GOs (e.g., visual display) 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.

Semantic mapping (SM) is a heuristic that enables students to recognize relevant information from lecture and text (e.g., main ideas, important supporting details) and organize that information for written or oral retell (Washington, 1988). In SM, students and/or the teacher create a visual representation of new or difficult vocabulary and any relationships existing among the different vocabulary (Bos & Anders, 1992). In addition, when teaching this type of GO, a teacher presents critical attributes of a concept along with examples and non-examples to help promote student discrimination and generalization (Deshler et al., 1996).

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).

Study Purpose

The purpose of this study was to replicate and extend the current evidence of the effectiveness of GOs with students with LD. While it has been hypothesized that visual displays will assist with maintenance and far transfer for students with LD (Mayer, 1979; Robinson et al., 1998; Robinson & Skinner, 1996; Vekiri, 2002), it has not been directly tested. Given that students with LD have difficulty with maintenance and far transfer (Baumann, 1984; Gajria et al., 2007; Holmes, 1985; Johnson, Graham, & Harris, 1997; Kim et al., 2004; Williams, 1993), it was important to attempt to validate this hypothesis. Specifically, this study addressed the following questions:

1. What is the effectiveness of a semantic mapping lesson compared to a semantic mapping lesson plus a visual display in improving factual recall during social studies verbal instruction for adolescent students with LD?
2. Does the addition of a visual display to the semantic mapping lesson improve maintenance effects for adolescent students with LD?
3. Does the addition of a visual display to the semantic mapping lesson improve far transfer effects for adolescent students with LD?

Study Hypotheses

Based on the review of theory and meta-analysis of studies of GOs, both the experimental (SM + visual display) and control group (SM only) of students with LD and low achieving students should demonstrate a large effect between pretest and posttest. However, students in the SM + visual display condition should perform significantly higher on tests of maintenance and far transfer.

Method

Experimental Design

A pretest-posttest comparison group design was used to investigate the effects of an SM lesson plus visual display versus an SM lesson alone on students' ability to gain and maintain factual knowledge from expository social studies material. In addition, a posttest only comparison group design was used to examine the effects of the SM lesson plus visual display versus an SM lesson alone on students' far-transfer ability. Stratified purposeful sampling was used to divide the students into three groups: (a) normally achieving; (b) students with LD; and (c) low-achieving. Once these groups were determined, students were randomly assigned to the treatment (SM + visual display) or control (SM only) groups.

Participants and Setting

The study took place in a rural school district in the eastern United States. Three eighth grade social studies inclusion classrooms were selected for the study based on their high density of students with LD. Out of a total of 76 students, parental and student informed consent was obtained for 62 students. Nineteen of these students were identified as having a specific learning disability in reading (e.g., basic skills, fluency, and/or comprehension), 36 students were normally-achieving, and seven students were selected by the classroom teacher as being low-achieving based on factors closely related to characteristics identified in the research literature examining low-achievers (Ford, 1996; McCoach & Siegle, 2001; Schunk, 1998). The demographics of the 62 participating students were similar to the district as a whole. Thirteen participants received free or reduced lunch, similar to the 24% of the entire district. Fifty-five of the students were Caucasian, four were African American, and three were Hispanic. Twenty-eight of the students were female and thirty-three were male. Across treatment and control conditions, independent *t*-tests demonstrated no significant differences between groups on demographics (see Table 1).

Table 1
Participant Characteristics by Treatment Conditions by Total

	Condition		
Characteristics	SM + Visual Display	SM Only	Total
Number of Participants	33	29	62
Normally-Achieving Students	19 m = 10, f = 9	17 m = 7, f = 10	36
Students with LD	10 m = 6, f = 4	9 m = 5, f = 4	19
Low-Achieving Students	4 m = 3, f = 1	3 m = 2, f = 1	7
Chronological Age			
M	170.20	167.46	168.92
SD	6.67	5.92	6.32
Economic Status			
Free/reduced lunch (*n*)	8	5	13
Race			
African American (*n*)	3	1	4
Caucasian (*n*)	29	26	55
Hispanic (*n)*	1	2	3

Note. SM = semantic mapping. Chronological age stated in months as of April 20, 2010.

Students with LD. Nineteen students were designated as having a primary, specific learning disability in reading. Each of these 19 students received their social studies instruction in the general education classroom. Across treatment and control conditions, independent *t*-tests demonstrated no significant differences between the groups of students with LD on their unique characteristics (see Table 2).

Teacher/Researcher. To ensure authenticity, all instruction was provided in the general education classroom at the normal time for each of the three classes. The primary researcher, with five years experience as a special and regular education teacher, served as the instructor for each of the classes. The classroom teacher remained in the room during each class period, but was situated behind and out of sight of the students.

Table 2
Students' with LD Characteristics by Treatment Conditions by Total

Characteristics	Condition: SM + Visual Display	SM Only	Total
Number of Students with LD	10	9	19
Sex			
Male	6	5	11
Female	4	4	8
IQ score (Full scale WISC-III)			
M	95.2	96.6	95.86
SD	8.3	9.1	8.68
Reading Achievement (WJ-III)			
M	78.7	80.3	79.46
SD	7.5	7.9	7.69
Pullout Service (*n*)	7	5	12
Comorbid Conditions			
MD (*n*)	2	1	3
AD/HD (*n*)	4	3	7
MD + AD/HD (*n*)	1	1	2

Note. Scores obtained from school files and were based on tests administered by school personnel within the previous four years. MD = mathematics disability, AD/HD = attention-deficit/hyperactivity disorder, SM = semantic mapping.

Materials

Prior to the study, in collaboration with a content expert (a state certified social studies teacher with a B.A. in history) and the classroom teacher (who had 29 years of eighth grade social studies teaching experience), the primary researcher selected *Feudalism in Middle Ages Europe* as the lesson topic for the study. This topic was derived from a ninth grade state standard, one year above the student level in this study. The ninth grade state standard was selected for content validity purposes (e.g., actual content the students are expected to learn in the future) and to mitigate the chances of prior knowledge affecting the study outcomes.

After selection of the topic, the social studies content expert created an expository passage to be used for instruction. The passage was 546 words long and fell at a 6.4 grade level based on a readability test. The Lexile level was 860L, which equals approximately a late sixth grade or early seventh grade reading level. The rationale to go below grade level was based on classroom practice. The classroom teacher reported most passages used for instruction over the course of the year fell in the 6 to 7.5 grade level range. Based on the expository passage, the primary researcher created an SM lesson wherein the instructor and students (both treatment and control groups) created a semantic map together. Following the suggestions of Gersten et al. (2005), the lesson was fully scripted to increase the likelihood of fidelity over the three class periods.

Prior to posttest, the treatment condition (SM + visual display) received a researcher-created visual display to study for 10 minutes. The control group (SM only) was only allowed to study the semantic map they created. The visual display provided to the treatment group is illustrated in Figure 1.

Overview of Instruction

The SM lesson was delivered based on the recommendations of Washington (1988) and included: (a) brief introduction; (b) questions and/or predictions; (c) vocabulary overview; (d) stated purpose; (e) reading the passage; (f) brainstorming; and (g) creating the map. Both treatment and control groups were taught concurrently and received the same amount of instruction. Each lesson lasted the fully allotted 45-minute class period.

Brief introduction. The brief introduction served as a connection between the students' background knowledge and the information they would be learning. For each of the three classes, the researcher introduced the concept of feudalism in Middle Ages Europe.

Questions and/or predictions. After the introduction, students were allowed to ask questions about anything they were curious about or make predictions about the passage. Across the three classes, no predictions were made. There was no potentially confounding student question or insight that might have given advantage to one class over another.

Vocabulary overview. Prior to the lesson, the social studies content expert and the classroom teacher identified 13 words from the passage they felt held significance for the lesson and might have been unfamiliar to the students. The vocabulary list included: feudalism, noble, peasant, knight, page, chivalry, squire, vassal, fief, manor, serf, moat, and waterwheel.

Stated purpose. Just prior to reading the passage, the researcher explicitly stated the purpose of the lesson. Specifically, the researcher stated to the class, "For this lesson, the purpose is to recall facts and ideas about feudalism in Middle Ages Europe. I want you to remember as many details from the passage as you can."

Reading the passage. Based on the knowledge that 12 students across the three classes received pullout instruction for passage reading and test taking, the researcher read the passage aloud to each of the classes, firming up main ideas and vocabulary after each paragraph. The common main ideas and common vocabulary were used to increase fidelity of implementation across the three classes.

Feudalism Hierarchy

Nobles
Kings and queens; only 10% of population
Lived in cold, drafty castles

↓

Knights
sons of nobles;
3 stages to become a Knight =
1. Page (learned chivalry),
2. Squire (learned to ride and fight),
3. Knight

↓

Vassals
lesser nobles;
granted fief (land) for promise to fight for the nobles

↓

Peasants
limited rights;
could operate private business

↓

Serfs
slaves; no rights

Figure 1. Visual Display

Brainstorming. After reading the passage to the students, the researcher asked the students to brainstorm any facts or ideas they recalled from the lesson. The researcher urged the students not to directly quote from the passage, but to put the ideas in their own words.

Creating the map. After the students ran out of ideas or facts to add to the list, the researcher used prompting to assist the students in creating the semantic map. A typical student-created semantic map from the lesson has been reproduced in Figure 2.

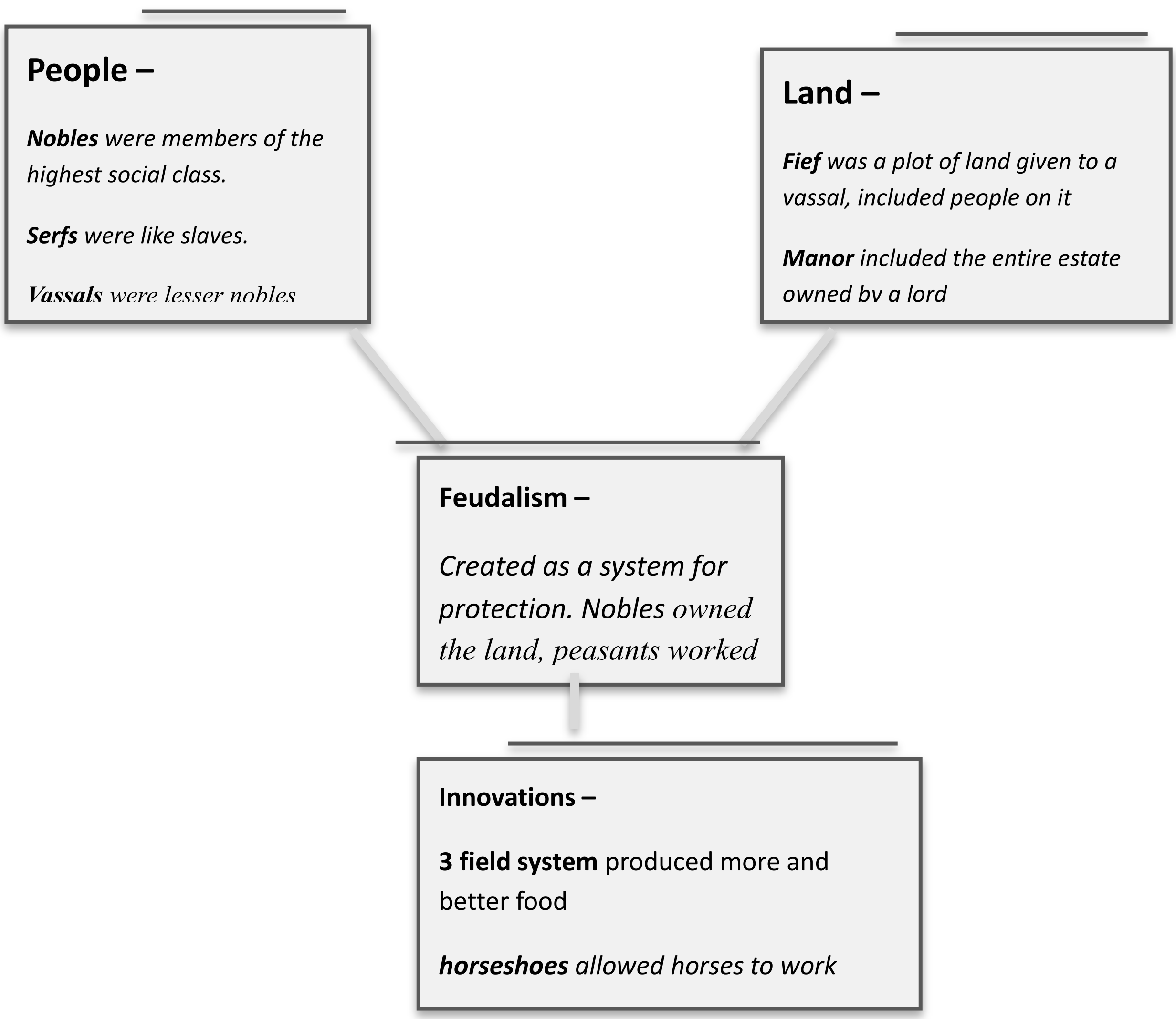

Figure 2. Completed Semantic Map

Fidelity of implementation. The three class sessions were audio recorded to verify instruction was delivered in the same manner for each class. Within the week following instruction, the primary researcher and a graduate research assistant separately analyzed the audio recordings with a copy of the lesson script. Special attention was given to adherence to the script, wording of definitions, and points emphasized. No major discrepancies in instruction were uncovered between classes. The major difference between each class was the student questions and responses. These varied somewhat by class, but it was determined that the researcher was able to successfully guide each class back to common anchoring points in adherence with the script.

Measures

Multiple measures were used at pretest, posttest, and maintenance to test the factual recall of all students in the study – a written factual recall measure and a multiple-choice measure. At each time of testing, the written factual recall measure preceded the multiple-choice measure so that the information in the multiple-choice test did not influence the written recall. Additionally, at posttest and maintenance, a five-question multiple-choice measure was added to test students' far-transfer ability.

Written factual recall. A written factual recall measure was used at pretest, posttest, and maintenance. Due to school district policy and request, the written factual recall measure was not a straight retell measure. The district literacy experts provided a five-paragraph essay outline worksheet for this measure. The worksheet provided a space for a main idea (i.e., thesis), three subordinate ideas, three details for each subordinate idea, and a conclusion statement. The students had much experience working with this worksheet and in other classes were encouraged to fill out the main and subordinate ideas before adding any details. As such, the students understood what to do when told to write in "five paragraph essay form."

The scoring of the written factual recall measure was based on the worksheet. One point was given for each reasonable and correct component (e.g., main idea, subordinate ideas, subordinate details) written in a complete or partial sentence. For example, in the case of an overall thesis statement, "Feudalism in Middle Ages Europe developed out of a need for protection and security" or "Feudalism about land and protection" earned one point; while "Feudalism contained lots of people" earned no points.

Multiple-choice. A multiple-choice measure containing 20 factual recall items was also used at pretest, posttest, and maintenance. This measure was developed by the social studies content expert based on the expository passage and the primary researcher was blind to the items on the test during creation of the scripted lesson. However, like the expository passage, the multiple-choice test was read aloud to the students. Because of this, the researcher was aware of the test items after pretest and before instruction, but the lesson script was not altered in any form. The same test was used at posttest and maintenance, but the questions and answer choices were randomly reordered before each subsequent testing.

The multiple-choice pretest was also used to test the internal consistency of the measure and control for prior knowledge. Because each of the items was dichotomously scored (i.e., 0 for incorrect, 1 for correct), the Kuder-Richardson 20 (KR-20) formula was used to determine internal consistency (i.e., how consistent subject responses are among the questions on an instrument). A reliability coefficient of .81 indicated individual test items produced similar patterns of responding across all participants. This confirmed the test items were homogenous and reliable for the pretest and alternate forms (i.e., posttest, maintenance).

To control for prior knowledge, any participant with more than 12 items correct at pretest would be excluded from the study. No participants were excluded for this reason (i.e., pretest range was 3 – 11 items correct).

Far transfer. Five multiple-choice questions were added to the posttest and maintenance test to measure students' ability to answer far-transfer items (e.g., similar relational content not covered in the lesson). For example, a sample question was "Similar to chivalry, bushido was the Japanese code of which group? (a) Geisha (b) Samurai (c) Rulers (d) Priests." The social studies content expert created the far-transfer items and they were interspersed with the 20 factual recall items.

Scoring and reliability. Initially, the primary researcher scored the written factual recall measures and multiple-choice tests. Afterwards, the researcher coded each measure from 1 – 62 and had two graduate research assistants score each measure for reliability purposes. The coding ensured the graduate students would be blind to condition and student name. For the written factual recall, 83% reliability was obtained initially between the three scorers. After discussion and reexamination among the scorers, the reliability increased to 95%. For the multiple-choice measures, reliability was 100%.

Social Validity. A student attitude measure allowed students to indicate how they felt about the instruction they received. Using a measure previously developed and tested (e.g., Darch & Carnine, 1986; Darch, Carnine, & Kameenui, 1986; Darch & Gersten, 1986), all subjects rated instruction, using a 5-point Likert scale, on three dimensions: how much they learned, whether they liked the SM lesson, and whether they liked studying with the visual display (treatment) or semantic map (control). This measure provided data on the social validity of the experiment.

Results

Written Factual Recall

Pretest-posttest. Descriptive and statistical data for the pretest and posttest written factual recall measure are displayed in Table 3. The *F*-statistic was a result of one-way analyses of variance (ANOVAs) used as significance testing between mean gain by condition.

Table 3
Pretest – Posttest Written Factual Recall

	Pretest		Posttest			
Condition	*M*	*SD*	*M*	*SD*	*Mean Gain*	*F*
SM + Visual Display *Overall, N* = 33	.51	.53	4.54	1.03	+4.03	10.01**
SM Only *Overall, N* = 29	.55	.69	3.64	1.27	+3.09	
Disaggregated by Student Type						
SM + Visual Display *Students with LD, N* = 10	.10	.32	4.10	.74	+4.00	25.59***
SM Only *Students with LD, N* = 9	.22	.44	2.89	.78	+2.67	
SM + Visual Display *Low Achieving Students, N* = 4	0	0	2.25	.50	+2.25	*n.s.*
SM Only *Low Achieving Students, N* = 3	0	0	1.67	.58	+1.67	
SM + Visual Display *Normally Achieving, N* = 19	.84	.76	5.26	1.66	+4.42	5.30*
SM Only *Normally Achieving, N* = 17	.82	.95	4.29	1.57	+3.47	

Note. *** = $p < .001$, ** = $p < .01$, * = $p < .05$

Overall, across student type and condition, students averaged less than one correct written statement at pretest. After disaggregating the data, it was shown that students with LD averaged only .16 correct written statements and low achieving students produced no correct written statements at pretest. Mean gains between pretest and posttest favored the SM + visual display group in all categories of students, but most significantly with students with LD, $F(1, 17) = 25.59$, $p < .001$.

Posttest and maintenance only. Results of the written factual recall measure were also analyzed for effect sizes (*ES*s) at posttest and maintenance. Effect sizes here and in subsequent analyses are reported as Cohen's *d* (> .2 = small effect, > .6 = moderate effect, > .8 = large effect; Cohen, 1988). In addition, one-way ANOVAs were used for significance testing. Overall results by condition are displayed in Table 4.

Table 4
Overall Written Factual Recall Posttest and Maintenance Only Effects by Condition

	SM +Visual Display		*SM Only*			
Measure	*M*	*SD*	*M*	*SD*	*ES*	*F*
Posttest	4.54	1.03	3.64	1.27	.78	9.28**
Maintenance	4.94	1.28	3.45	1.32	1.15	18.03***
	N = 33		*N* = 29			

Note. ES = Effect size. Both effect sizes in favor of the SM + visual display group.
***$p < .001$, **$p < .01$

A moderate *ES* favoring the SM + visual display condition was found for posttest, while a strong *ES* was found for maintenance. These effects were both statistically significant. Of note, the mean number of correct written factual statements increased between posttest and maintenance for the SM + visual display condition, while it decreased for the SM only condition. Thus, a larger effect for maintenance was demonstrated across all students. These data were also disaggregated by student type. The results are displayed in Table 5.

Effects favored the SM + visual display group across each student type for posttest and maintenance. Students with LD demonstrated the largest effects for posttest and maintenance, both strong (e.g., > .8) and statistically significant. Low achieving students displayed large effects, but due to such small sample sizes the effects were not statistically significant. The normally achieving group demonstrated a strong, statistically significant effect for maintenance only. In terms of correct written factual statements, students with LD were the only group whose mean number decreased between posttest and maintenance. The low achieving and normally achieving groups both saw an increase between posttest and maintenance for the SM + visual display condition.

Table 5

Disaggregated Written Factual Recall Posttest and Maintenance Only Effects by Condition

	SM +Visual Display		SM Only			
Group/Measure	*M*	*SD*	*M*	*SD*	*ES*	*F*
Students with LD						
Posttest	4.10	.74	2.89	.78	1.59	11.78**
Maintenance	4.00	.67	2.67	1.12	1.46	9.85*
	N = 10		*N* = 9			
Low Achieving Students						
Posttest	2.25	.50	1.67	.58	1.09	*n.s.*
Maintenance	2.50	.58	2.33	.58	1.29	*n.s.*
	N = 4		*N* = 3			
Normally Achieving Students						
Posttest	5.26	1.66	4.29	1.57	.60	*n.s.*
Maintenance	5.95	1.75	4.06	1.56	1.14	11.95**
	N =19		*N* = 17			

Note. ES = Effect size. All effect sizes in favor of the SM + visual display group.
**$p < .01$, * $p = < .05$

Multiple-Choice Factual Recall

Pretest – posttest. Descriptive and statistical data for the pretest and posttest multiple-choice factual recall measure are displayed in Table 6. The *F*-statistic was a result of one-way ANOVAs used as significance testing between mean gain by condition.

Table 6
Pretest-Posttest Multiple-Choice Factual Recall

	Pretest		Posttest			
Condition	*M*	*SD*	*M*	*SD*	*Mean Gain*	*F*
SM + Visual Display	7.24	2.44	14.85	3.05	+7.61	
Overall, N = 33						4.31*
SM Only	7.55	2.15	13.76	2.91	+6.21	
Overall, N = 29						
Disaggregated by Student Type						
SM + Visual Display	6.70	1.89	13.80	1.81	+7.10	
Students with LD, N = 10						*n.s.*
SM Only	6.56	1.74	13.44	1.81	+6.88	
Students with LD, N = 9						
SM + Visual Display	5.25	1.71	11.00	1.83	+5.75	
Low Achieving Students, N = 4						*n.s.*
SM Only	6.00	2.00	9.67	3.06	+3.67	
Low Achieving Students, N = 3						
SM + Visual Display	7.95	2.59	16.21	2.92	+8.26	
Normally Achieving, N = 19						5.01*
SM Only	8.35	2.09	14.65	2.83	+6.30	
Normally Achieving, N = 17						

Note. * = $p < .05$

Overall, across student type and condition, students averaged 7.4 correct answers (out of 20) at pretest on the multiple-choice factual recall measure. The average increased to 14.31 correct answers (out of 20) at posttest. Across all students, there was a significant difference in mean gain, $F(1, 60) = 4.31$, $p < .05$, favoring the SM + visual display group. After disaggregating the data, mean gains between pretest and posttest favored the SM + visual display group in all categories of students, but most significantly with the normally achieving group, $F(1, 34) = 5.01$, $p < .05$.

Posttest and maintenance only. Results of the multiple-choice factual recall measure were also analyzed for *ES*s at posttest and maintenance. In addition, one-way ANOVAs were used for significance testing. Overall results by condition are displayed in Table 7. The means are out of a total of 20 possible items.

Table 7
Overall Multiple-Choice Posttest and Maintenance Only Effects by Condition

	SM +Visual Display		*SM Only*			
Measure	*M*	*SD*	*M*	*SD*	*ES*	*F*
Posttest	14.85	3.05	13.76	2.91	.33	*n.s.*
Maintenance	13.55	3.25	11.14	2.81	.78	9.62**
	N = 33		*N* = 29			

Note. ES = Effect size. Both effect sizes in favor of the SM + visual display group.
**$p < .01$

A small *ES* favoring the SM + visual display condition was found for posttest, although it was not statistically significant. However, a significant moderate effect was found for maintenance across all students, favoring the SM + visual display condition. The mean correct number of multiple-choice items between posttest and maintenance decreased by 1.3 in the SM + visual display group, while the decrease was 2.62 items for the SM only group. These data were also disaggregated by student type. The results are displayed in Table 8.

Effects favored the SM + visual display group across each student type for posttest and maintenance. For posttest, all effects were small. However, for maintenance, all effects were strong and statistically significant for students with LD and for normally achieving students. Students with LD demonstrated the largest effects for maintenance, a robust 1.41. Low achieving students displayed a large effect for maintenance, but it was not statistically significant. Of particular note, the students with LD in the SM + visual display group only decreased by .4 items correct between posttest and maintenance, while the SM only group decreased by 2.44 items correct. The students with LD in the SM + visual display condition had the highest level of maintenance in relation to posttest scores of the three groups of students.

Table 8
Disaggregated Multiple-Choice Posttest and Maintenance Only Effects by Condition

	SM +Visual Display		SM Only			
Group/Measure	*M*	*SD*	*M*	*SD*	*ES*	*F*
Students with LD						
Posttest	13.80	1.81	13.44	1.81	.22	*n.s.*
Maintenance	13.40	1.78	11.00	1.58	1.41	9.57**
	N = 10		*N* = 9			
Low Achieving Students						
Posttest	11.00	1.83	9.67	3.06	.55	*n.s.*
Maintenance	9.50	3.70	6.67	2.08	.89	*n.s.*
	N = 4		*N* = 3			
Normally Achieving Students						
Posttest	16.21	2.92	14.65	2.83	.56	*n.s.*
Maintenance	14.47	3.22	12.00	2.74	.84	6.07*
	N =19		*N* = 17			

Note. ES = Effect size. All effect sizes in favor of the SM + visual display group.
**$p < .01$, * p = < .05

Far Transfer

Posttest and maintenance only. Results of the multiple-choice far transfer measure were analyzed for *ES*s at posttest and maintenance. In addition, one-way ANOVAs were used for significance testing. Overall results by condition are displayed in Table 9. The means are out of a total of five possible items.

Table 9

Overall Far-Transfer Effect by Condition

	SM +Visual Display		SM Only			
Measure	*M*	*SD*	*M*	*SD*	*ES*	*F*
Far Transfer	3.33	1.11	2.72	1.19	.53	4.36*
Maintenance	3.21	.89	2.31	1.26	.83	16.80***
	N = 33		*N* = 29			

Note. ES = Effect size. Both effect sizes in favor of the SM + visual display group.
*** = $p < .001$, * = $p < .05$

A statistically significant moderate *ES* favoring the SM + visual display condition was found for posttest. In addition, a significant strong effect was found for maintenance across all students, also favoring the SM + visual display condition. The difference between mean correct numbers of far transfer items between conditions was .61 at posttest and increased to .9 at maintenance. These data were also disaggregated by student type. The results are displayed in Table 10.

Effects favored the SM + visual display group across each student type for far-transfer posttest and maintenance. For far-transfer posttest, there was a strong effect for students with LD and the low achieving group, although only the *ES* for students with LD was statistically significant. The normally achieving group had only a small effect for posttest. For far-transfer maintenance, both students with LD and low achieving students had a strong, statistically significant effect. The normally achieving group had only a small maintenance effect. Of particular note, the low achieving group in the SM + visual display condition increased by .25 items correct between posttest and maintenance, while the SM only group decreased by .66 items correct. For both students with LD and normally achieving students in the SM + visual display condition, far -transfer results were almost identical for posttest and maintenance.

Table 10

Disaggregated Far Transfer Effect by Condition

	SM +Visual Display		SM Only			
Group/Measure	*M*	*SD*	*M*	*SD*	*ES*	*F*
Students with LD						
Far Transfer	3.51	.85	2.11	.78	1.70	13.88**
Maintenance	3.32	.67	1.78	.97	1.84	16.51***
	N = 10		*N* = 9			
Low Achieving Students						
Far Transfer	2.00	.82	1.33	.58	.91	*n.s.*
Maintenance	2.25	.50	.67	.58	2.96	15.04**
	N = 4		*N* = 3			
Normally Achieving Students						
Far Transfer	3.53	1.12	3.29	1.10	.21	*n.s.*
Maintenance	3.37	.96	2.88	1.11	.47	*n.s.*
	N =19		*N* = 17			

Note. ES = Effect size. All effect sizes in favor of the SM + visual display group.
*** = $p < .001$, ** = $p < .01$

Social Validity

Students were asked to answer three questions on a 5-point Likert scale (e.g., 1 = very little, 5 = very much). The results of the measure were separated by condition. For each question, mean by condition is displayed in Table 11.

Overall, students in both experimental conditions reported they felt they learned a lot and enjoyed the SM lesson, as well as the opportunity to study with either the visual display or semantic map. There were no statistically significant differences between conditions on any of the three questions. After disaggregating the data, the students with LD, regardless of condition, supplied the highest ratings for each question. For example, for question 1, the means were 4.6 and 4.56 for the SM + visual display and SM only conditions, respectively. Question 2 yielded means of 4.9 and 4.89 and question 3 yielded means of 4.8 and 4.67, respectively.

Table 11
Social Validity by Condition

	SM +Visual Display	SM Only
Question	*M*	*M*
1. How much do feel you learned?	4.42	4.48
2. Did you like the semantic mapping lesson?	4.79	4.69
3. Did you like studying with the visual display (treatment) or semantic map (control)?	4.64	4.55
	N = 33	N = 29

Discussion

The results of this study supported the conclusion that semantic mapping was beneficial for factual recall, while the additive effect of a visual display significantly improved maintenance and far transfer for adolescents with LD. Results of this study also supported the conclusion that normally achieving students and low achieving students also benefit from semantic mapping and the visual display. This finding was consistent over written and multiple-choice measures.

Written Factual Recall

The written factual recall measure tested the students' ability to produce newly acquired knowledge in essay form. Unfortunately, due to the school district request to use the five-paragraph outline worksheet, the measure ultimately tested only isolated facts the students could remember and write in sentence form. Making matters worse, the time limitations (i.e., seven minutes) precluded any chance at depth or inferential/relational statements as students had very little time to brainstorm and plan their writing effort.

However, these issues aside, the data extracted from the written measures were consistent with the multiple-choice data, and yielded enough information for analysis.

Between pretest and posttest, the SM + visual display condition had a significantly larger mean gain increase in factual statements compared to the SM only condition (4.03 compared to 3.09). An analysis of posttest only effects also favored the SM + visual display condition, with a moderate significant effect ($ES = .78$, $p < .05$). After these data were disaggregated by student type, it was evident that students with LD had the largest mean gain increase between conditions (4.00 compared to 2.67) and the only significant effect when comparing posttest only ($ES = 1.59$, $p < .01$), both favoring the SM + visual display condition. All other students (i.e., normally achieving and low achieving) made large gains from pretest to posttest on the written measures, but no significant differences were found between experimental conditions at posttest. This finding was somewhat surprising. Our hypothesis was that semantic mapping on its own would drive initial acquisition. Therefore, we did not expect such a strong and significant effect for the SM + visual display condition by students with LD at posttest. It is clear the additive effect of the visual display assisted students with LD beyond semantic mapping on its own for writing factual statements at posttest.

When the written factual recall measure was administered again for maintenance, 10 days after posttest, an overall significant strong effect ($ES = 1.15$, $p < .01$) was found favoring the SM + visual display condition. However, unlike posttest, there were large significant maintenance effects for both students with LD ($ES = 1.46$, $p < .05$) and normally achieving students ($ES = 1.14$, $p < .01$) after disaggregating by student type. Large effects were also found for low achieving students, but sample limitations (sample size = 7) negate their statistical significance. The maintenance results clearly match our hypothesis and support our previous finding that the additive effect of the visual display in addition to the semantic map helps students retain newly acquired factual knowledge (Dexter & Hughes, 2010).

The results of the written factual recall measure were consistent with the results of previous research examining GOs and recall of ideas and details in writing for students with LD (e.g., Draheim, 1983; Ruddell & Boyle, 1989; Sturm & Rankin-Erickson, 2002). Each of those studies also found students with LD were able to recall more factual details after attending to a GO. However, each of those studies took place in a resource room setting after regular school hours. The results presented in this study extend the literature on written factual recall by utilizing an inclusion classroom during the regular school day. In this natural setting, students with LD improved significantly, as did their normally achieving peers. Unfortunately, due to the brevity of this study, written measures accounting for relational or inferential statements, increased length, and improved holistic scores could not be administered. By limiting the results of the written measure to only factual recall, some important information often associated with GO research (e.g., attaining relational knowledge) was sacrificed (DiCecco & Gleason, 2002).

Multiple-Choice Factual Recall

Like the majority of studies on GOs with students with LD, this study measured factual recall using a multiple-choice test (Dexter & Hughes, 2010; Gajria et al., 2007; Kim et al., 2004). However, as the previous reviewed studies typically occurred in resource room or after school settings, this study utilized an inclusion classroom during regular school hours. This extends the previous research by testing effects in a more naturally occurring school environment with many types of students included.

As was hypothesized, even though there was a significant overall mean gain increase between pretest and posttest favoring the SM + visual display group (e.g., 7.61 compared to 6.21), there was no significant effect by condition at posttest only overall, or after disaggregation by student type. This supports our previous finding that semantic mapping by itself is effective for initial acquisition (Dexter & Hughes, 2010). Furthermore, the additive effect of the visual display was seen in the maintenance results ten days after posttest. Overall, the SM + visual display condition significantly outperformed the SM only condition at maintenance (*ES* = .78, $p < .01$). After disaggregating the results, large significant effects were found for students with LD (*ES* = 1.41, $p < .01$) and normally achieving students (*ES* = .84, $p < .05$). A large effect was also found for the low achieving group (*ES* = .89), but it did not reach statistical significance due to the small sample size ($N = 7$). Like the results of the posttest, this confirms our hypothesis and supports our previous finding that the additive effect of the visual display in addition to the semantic map is crucial for retention of newly acquired factual knowledge (Dexter & Hughes).

While these results confirmed our hypothesis and were promising, it is important to point out that even for the top overall student condition (SM + visual display) the mean multiple-choice posttest score was 14.85 out of 20. This equals 74.25% accuracy. While this would not be considered ideal by any teacher's standard, it is based on one class period of instruction on new material followed by a delayed (e.g., next day) posttest. This limitation should be addressed in future research.

Far Transfer

This study also measured students' far-transfer ability (i.e., applying knowledge to situations not directly covered in the text or lecture) using a multiple-choice measure. Previous reviews of GO research with students with LD (e.g., Dexter & Hughes, 2010; Gajria et al., 2007; Kim et al., 2004) indicate that GOs may improve inference skills and relational knowledge for secondary students with LD. However, the evidence is limited due to the few studies explicitly measuring far transfer (Dexter & Hughes).

Across all students, there was an overall moderate far-transfer effect favoring the SM + visual display group at posttest (*ES* = .53, $p < .05$). As was hypothesized, after disaggregation, a large significant effect was found for students with LD (*ES* = 1.70, $p < .01$), while only a small effect (*ES* = .21) was found for the normally achieving group. A large effect (*ES* = .91) was found for low achieving students, but again, did not reach statistical significance due to sample size. This finding is consistent with Mayer's (1979) contention that GOs assimilate material to a broader set of past experiences allowing students with lower verbal ability to more successfully transfer verbal information to new situations, while it may not be necessary for students with higher verbal ability (e.g., normally achieving students).

Likewise, at maintenance, students with LD demonstrated a significant large effect ($ES = 1.84$, $p < .001$) for the SM + visual display group, while the normally achieving group demonstrated only a moderate effect ($ES = .47$). Furthermore, the low achieving group, despite the small sample size, reached a statistically significant large effect at maintenance ($ES = 2.96$, $p < .01$). This supports our hypothesis and previous finding that the additive effect of a visual display to a semantic mapping lesson may bridge the connection of verbal information with prior knowledge and assist students with low verbal ability in far-transfer tasks over longer periods of time (Dexter & Hughes, 2010).

Social Validity

This study also measured students' attitude toward the semantic mapping lesson, the GO they used to study before posttest, and how much they felt they learned. Across all students, the mean scores indicate students liked "very much" the semantic mapping lesson and the GO they used to study, regardless of type. All students also perceived they learned a lot from the lesson. In addition, the classroom teacher was impressed with the results and reported he will use this type of lesson and study format in the future.

Limitations and Directions for Future Research

While the results of this study are promising, there are two significant limitations to this research. First, the measures used in this study primarily focused on factual recall and far transfer. Focusing on these outcomes limited our ability to measure relational and inferential knowledge, which are important for GO research (DiCecco & Gleason, 2002). It is also important to note that all measures were created by a social studies expert and closely tied to the content. While the included measures should have good content validity, there is no way to measure broader construct validity. This fact may limit the generalizability of these findings, and questions the actual level of understanding obtained by students across conditions (Boyle, 1996). Future research should find ways to include relational and inferential measures. Oral retell is a measure that could potentially assess factual recall, as well as more relational or inferential statements. Also, where possible, standardized measures could be used to measure broader construct validity.

The second significant limitation to this research was its brevity. There was only one day of instruction with the semantic mapping lesson that was new to all students. Previous research with GOs suggests a timeframe of four to six weeks for successfully implementing a GO intervention program (DiCecco & Gleason, 2002; Gajria et al., 2007). The positive effects for this study under such a short duration are promising. Future research should seek to test this kind of GO program over a longer period of time. Consistent use of these types of

GOs over time will produce more far-reaching results and better inform inclusionary practice.

Implications for Practice

Consistent with the findings from the meta-analysis, this study found that an instructionally intensive type of GO (e.g., semantic mapping) worked well for immediate factual recall across conditions, while the addition of a more computationally efficient GO (e.g., visual display) produced larger maintenance and transfer effects than semantic mapping alone. These results can help teachers in designing GOs for initial instruction and for re-teaching, studying, and retention purposes. As in this study, 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. The retention aspect has special relevance to secondary students with LD who must be able to retain knowledge learned in school for statewide testing and promotion/graduation purposes.

Additionally, this study found that effects went beyond students with LD to low achieving students and normally achieving students. All students improved significantly between pretest and posttest on factual recall measures. All students, regardless of type, also demonstrated at least a small effect on posttest and maintenance only measures, as well as far-transfer measures. There were no negative effects across any condition or any type of student. This finding lends support to the benefits of GOs for inclusive classrooms. Furthermore, this study found all students enjoyed using the GOs and felt they learned a great deal.

The evidence in this study should persuade educational practitioners to make well-planned and well-instructed use of graphic organizers. A thoughtful combination of types of graphic organizers will help make the learning process more efficient for all secondary students, especially those students with LD.

References

Ausubel, D.P. (1960). The use of advanced organizers in learning and retention of meaningful material. *Journal of Educational Psychology, 51*, 267-272.

Ausubel, D. P. (1968). *Educational psychology: A cognitive view.* New York: 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.

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

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., & 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.

Dexter, D. D., & Hughes, C.A. (2010, April). *Graphic organizers and students with LD: How theory guides practice.* Paper presented at the International Council for Exceptional Children Conference, Nashville, TN.

Dexter, D.D., & Hughes, C.A. (2011). Graphic organizers and students with learning disabilities: A meta-analysis. *Learning Disability Quarterly, 34*, 51- 72.

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

Draheim, M. E. (1983). Facilitating comprehension and written recall of exposition through DRTA instruction and conceptual mapping. *Paper presented at the Annual Meeting of the National Reading Conference*. Austin, TX.

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

Ford, D. Y. (1996). Determinants of underachievement as perceived by gifted, above average, and average black students. *Roeper Review, 14*, 130- 136.

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.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M.S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children, 71*, 149- 164.

Graham, S., & Hebert, M.A. (2010). *Writing to read: Evidence for how writing can improve reading. A Carnegie Corporation Time to Act Report.* Washington, DC: Alliance for Excellent Education.

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.

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.

Larkin, J. H., & Simon, H.A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65 – 99.

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

McCoach, D. B., & Siegle, D. (2001). A comparison of high achievers' and low achievers' attitudes, perceptions, and motivations. *Academic Exchange, 2*, 71 – 76.

McCrudden, M. T., Schraw, G., Lehman, S., & Poliquin, A. (2007). The effect of causal diagrams on text learning. *Contemporary Educational Psychology, 32*, 367 – 388.

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

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

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

Ruddell, R. B., & Boyle, O. F. (1989). A study of cognitive mapping as a means to improve summarization and comprehension of expository text. *Reading Research and Instruction, 29*, 12- 22.

Sadoski, M., & Paivio, A. (2001). *Imagery and text: A dual coding theory of reading and writing.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Schunk, D.H. (1998, November). *Motivation and self-regulation among gifted learners.* Paper presented at the annual meeting of the National Association of Gifted Children, Louisville, KY.

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, L., & Kim, K. (2007). Working memory, short-term memory, and naming speed as predictors of children's mathematical performance. *Intelligence, 35*, 151 – 168.

Vekiri, I. (2002). What is the value of graphical displays in learning? *Educational Psychology Review, 14*, 261 – 313.

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

Washington, V. M. (1988). Semantic mapping: A heuristic for helping learning disabled students write reports. *Reading, Writing, and Learning Disabilities, 4*, 17- 25.

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.