Improving First-Year Writing Using Argument Diagramming

Maralee Harrell (mharrell@andrew.cmu.edu)
Carnegie Mellon University, Department of Philosophy, 5000 Forbes Ave.
Pittsburgh, PA 15213 USA

Danielle Wetzel (dfz@andrew.cmu.edu)
Carnegie Mellon University, Department of English, 5000 Forbes Ave.
Pittsburgh, PA 15213 USA

Abstract

There is substantial evidence from many domains that visual representations aid various forms of cognition. We aimed to determine whether learning to construct visual representations of argument structure enhanced the acquisition and development of argumentative writing skills within the context of first-year college writing course. We found a significant effect of the use of argument diagrams, and this effect was stable even when multiple plausible correlates were controlled for. These results suggest that natural—and relatively minor—modifications to standard first-year composition courses could provide substantial increases in student writing ability.

Keywords: argument diagramming; argument mapping; writing; critical thinking; graphic organizers.

Introduction

The purpose of the First-Year Writing (FW) Program at Carnegie Mellon University is to develop the academic reading and writing skills each student needs to be successful in his or her college career. Each student at CMU must take the course *Interpretation and Argument*, which is the core of this writing program.

Thus, though not titled 'Critical Thinking,' the FW course taken during the first year is generally one of the student's first introductions to thinking critically at a college level. Among other goals, the specific learning objectives for the FW Program is for students to be able to: (a) analyze a written argument by identifying the conclusion and the premises (both implicit and explicit) and describe how the premises support the conclusion, (b) evaluate a written argument by determining whether the premises do in fact support the conclusion, and whether the premises are reasonable, and (c) write an essay that both analyzes and evaluates one or more arguments.

The over-arching goal for the FW course is to provide foundational reading and writing skills that will enable students to develop advanced literacy in their own disciplines.

Most educators agree that one aspect of "critical thinking" involves the ability to reconstruct, understand and evaluate an argument—cognitive tasks we may describe as 'argument analysis' (see, e.g., Ennis, 1987; Fisher & Scriven, 1997; Kuhn, 1991). In college, the most common medium through which arguments are analyzed is writing. *Interpretation and Argument* is a research-based course that understands that reading and writing are inseparable practices for college-level course work. In the course, students are exposed to a variety of different texts (mostly academic essays) so they can explore a single issue from multiple perspectives and eventually contribute an argument of their own to the discussion. Both the exploration and the contribution rely heavily on argument analysis at various stages.

The first step in this analysis is reading a text for the argument, as opposed to, for example, reading for the plot (as in a novel) or for the facts (as in a textbook). Mandler (1984) provides an overview of research supporting the claim that adults and children as young as 3-years-old possess "story schemata" that guide understanding when reading or listening to a story. Thus, learning the skill of reading for the argument requires students to develop a new schema, or set of schemata, with which they can interpret the text appropriately.

Schema theory, first introduced by Bartlett (1932, 1958) and further developed by Evans (1967), Mandler (1984) and Rumelhart and Ortony (1977), explains cognition as information processing mediated by schemata. A schema is a packet of knowledge containing both data and information about the interconnections among the data. Rumelhart (1980) refers to schemata as the representations of concepts stored in memory, and Sweller (1994) describes schemata as representations of either concepts or problem-solution procedures.

To facilitate the acquisition of new schemas, Sweller (1994) recommends reducing the extraneous cognitive load during the learning process. One common way of reducing extraneous cognitive load is by using graphic organizers (GOs), such as diagrams, to supplement regular reading and instruction. Previous research has shown that students' use of GOs is generally efficacious in producing improvements on a wide range of cognitive tasks — including those generally labelled CT tasks — that are significantly higher than improvements gained by students engaged in reading and regular instruction alone (Horton, et al., 1993; Moore & Readance, 1984). Thus, we are particularly interested in the efficacy of alternative teaching methods that incorporate GOs to increase argumentative writing performance.

In what might these alternative methods consist? Both Larkin and Simon (1987) and Winn (1991) argue that diagrammatic representations of information can make recognition of important features and drawing inferences

easier than a sentential representation of the same information. Indeed, research on student learning has consistently shown the efficacy of using diagrams to aid text comprehension (Armbruster & Anderson, 1984; Dansereau, et al.; Novak & Gowin, 1984; Schwartz & Rafael, 1985), as well as vocabulary development, postreading activities and writing preparation (Johnson, et al., 1986).

One candidate alternative teaching method, then, is instruction in the use of argument diagrams as an aid to argument comprehension and evaluation (see Figure 1).

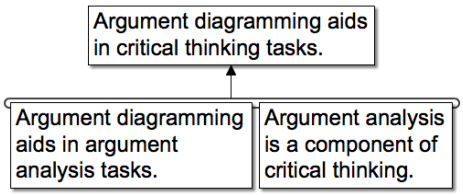

Figure 1: Example of a diagram for a simple argument.

If we think of an argument the way that philosophers and logicians do—as a series of statements in which one is the conclusion, and the others are premises supporting this conclusion—then an argument diagram is a visual representation of these statements and the inferential connections between them.

How does argument diagramming develop new schema? The argument diagramming curriculum consists in an online course introducing argument diagramming, followed by in-class and weekly homework assignments on representing the arguments in the course materials in diagrams. The students received oral and written feedback on their diagramming. The students are taught to discriminate between statements (or claims) and other kinds of sentences, as well as the difference between arguments and explanations. The students are also taught to look for words that indicate conclusions (e.g., 'thus' and 'therefore'), premises (e.g., 'because' and 'since'), linked arguments (e.g., 'but' and 'since') and convergent arguments (e.g., lists). All of these types of exercises help students develop an 'argument schema' for reading arguments in a variety of genres.

Recent research on the efficacy of an argument diagramming curriculum on the development of critical thinking skills includes studies on both philosophy students in introductory classes and a mix of undergraduates in critical thinking and informal logic classes. The former studies have shown that instruction that includes the use of argument diagrams to analyze, evaluate and create arguments significantly improves students' critical thinking skills over the course of a semester (Harrell, 2008, 2011, 2012).

The latter studies specifically on *computer-supported* argument visualization have shown that the use of software specifically designed to help students construct argument diagrams significantly improves critical thinking abilities over the course of a semester (Kirschner, Shum, and Carr 2003; Twardy 2004; van Gelder, Bissett, & Cumming, 2004). Additionally, research in this area has shown that student's critical thinking about specific topics is improved if students collaborate on argument diagram instruction instead of working alone (Scheuer, McLaren, Harrell, & Weinberger, 2011). This previous research, however, has all focused on performance on critical thinking skills tests—especially multiple choice tests like the California Critical Think Skills Test—and not on *writing tasks*.

Even so, we conjectured, that incorporating argument diagramming into our standard curriculum in *Interpretation and Argument* would help students develop their argumentative writing skills.

Hypothesis: Students who are able to construct argument diagrams and use them during argument analysis tasks will improve in performance on argumentative writing tasks over the course of a semester long composition class significantly more than students in the same class who do not have this ability.

Our first-year writing course was a natural place to study the skills acquisition of our students. We typically teach 28-30 sections of this course each semester, with a different instructor for each section. While the general curriculum of the course is set, including the sequence of assignments, each instructor is free to choose the readings for his or her section. The students who take this course are a mix of all majors from each of the seven colleges across the University. This study tests this hypothesis by comparing the pretest and posttest scores of students in Interpretation and Argument who were taught argument diagramming to the scores of those students who were not during the Fall of 2009, and the Spring and Fall of 2010.

Method

Participants

Eighty-one students (39 women, 42 men) across 7 sections of Interpretation and Argument were studied. In each semester, each section of the course had a different instructor and the students chose their section. Over the three semesters there were 7 different instructors. The students taught by Instructors 2, 6 and 7 were taught the use of argument diagrams to analyze the arguments in the course readings, while the students in the other sections were taught more traditional methods of analyzing arguments.

Materials and Procedure

We developed a pretest to be taken at the beginning of the semester, and a companion posttest to be taken at the end. For the next three semesters, students in both the treatment and control groups completed the pretests during the first week of the semester, and the posttest during the last week of the semester. Each test consisted in reading some text and completing two tasks. In Task 1, the student was asked to write an essay analyzing the argument presented by the author in the text. This analysis was to consist in identifying both the content and the structure of the argument. In Task

2, the student was asked to write an essay evaluating the argument presented by the author in the same text. The evaluation was to consist in a claim about the quality of the argument, and reasons to support that claim.

Results

Salient Features of Students' Writing

We recognize that text features alone do not constitute "good writing" and that there is no "right way" to read or write a text. We also recognize that privileging some text features over others might ignore other significant features. The features that we chose will help us locate change in demonstrable critical thinking between the pretest and posttest. We analyzed the texts for markers of text development and text coherence. We were interested in seeing to what extent there would be any kind of change in how many different ideas students could generate—about someone else's argument and about their own arguments. Within this category of "development," we identified the following for both Tasks 1 and 2 of the pre- and posttests: the number of *different* reasons or premises offered for the argument conclusion, and the number of *counterarguments* considered within the text.

For Task 1, we wanted to determine how much the students were understanding the argument in the text and what statements they would prioritize in their representations of it. For Task 2 only, we also considered whether students provided evidence or elaboration of their reasons. We wanted to distinguish between reasons that were supported with evidence and those that were not. Our concern was instances when students produced a lot of different ideas but failed to support them; we did not want to report "growth" in development without attempting to represent to what extent students were actually supporting their claims.

Because the number of ideas alone does not necessarily equate with good writing, and, in fact, one could argue that too many different ideas within an argument will result in chaos for a reader, we also looked for features that signaled an overall coherence in a written text. Vande Kopple has defined coherence as "prose in which nearly all the sentences have meaningful connections to sentences that appear both before and after them" (1989, 2). We also draw upon Enkvist's definition of coherence, "the quality that makes a text conform to a consistent world picture and is therefore summarizable and interpretable" (1990, 49). So, by *coherence*, we mean those features that enable a reader to make particular kinds of connections within the text. In coding Task 1, we considered the following as coherence markers: logical connections between premises and the argument conclusion, and logical connections between different premises

In coding Task 2, we looked at the following as markers of coherence: logical connections between premises and the argument conclusion, logical connections between different premises, and metacommentary (or "metadiscourse"). *Metacommentary* is language that writers use, according to Hyland (2003), to compose a text that is clear to a reader.

By providing linguistic "signposts" to readers, writers can create the effect that a text is coherent and holds together in an intentional way. Because these bits of language give clues for making sense of the text, their presence in a text can indicate that a writer is aware of a reader's needs for navigating the text successfully. These bits of language can also show that a writer understands his or her own text in particular ways and can point to a writer's strategic view of his or her writing. We were only interested in the effect that metacommentary has upon the readers—we were not interested in counting the different types. Therefore, coders scored Task 2 holistically for effective use of metacommentary.

Test Coding

Pretests and posttests were paired by student, and single-test students were excluded from the sample, resulting in 81 pairs of tests. The tests were coded during one extended session, using one set of coders for Task 1, and a different set for Task 2. Each coder independently coded all pairs of tests in his or her group (162 total tests). Each pre-/post-test pair was assigned a unique ID, and the original tests were blinded. To ensure reliability and validity, prior to each coding session, we had an initial coding-calibration session in which we and the coders coded several of the unpaired tests, discussed the codes, and came to a consensus about each code. After this, each coder was given the tests to be coded in a unique random order.

The categories to be coded for Task 1 were: Argument Conclusion, Counter-arguments, Premises, Connections and Errors. "Argument Conclusion" received a code of 1 if the student identified the conclusion of the argument, and a code of 0 if not. "Counter-arguments" received a code that indicated how many counter-arguments the student identified in the text. "Premises" received a code that indicated how many premises the student identified in the text. "Connections" received a code that indicated how many connections between premises or between a premise and the conclusion the student identified in the text. Finally, "Errors" received a code that indicated how many errors the student made; errors identified by the coders were (a) misunderstands counter-argument, (b) missing a major concept, (c) misreading (e.g. overstatement with no qualifiers), (d) misapplied quotation that shows disconnected reading, and (e) other.

The categories to be coded for Task 2 were: Conclusion, Premise, Evidence, Mismatch, Connections, Counter-arguments, and Metacommentary. "Conclusion" received a code of 1 if the student stated a thesis, and a code of 0 if not. "Premises" received a code that indicated how many premises the student used in support of the thesis. "Evidence" received a code that indicated how many premises were supported by evidence. "Mismatch" received a code that indicated whether, for each premise, the evidence offered actually supported that premise.

"Connections" received a code that indicated how many connections between premises or between a premise and the conclusion the student identified in the text. "Counter-arguments" received a code indicating how many counter-arguments the student considered. Finally, "Metacommentary" received a code of 0 if there was no metacommentary, 1 if the metacommentary was present but weak, and 2 if the metacommentary was strong. Then, for each task, the codes from the two coders on these categories were averaged, allowing for a more nuanced scoring of each category than either coder alone could give.

For each task, the primary variables of interest were the individual averages for each category on the pretest and the posttest. In addition, however, the following data was recorded for each student: the student's math, writing and verbal scores on the SAT, the section in which the student was enrolled, the student's final grade in the course, the student's home college, the student's sex, and whether the student had been taught using the AD curriculum.

Student Characteristics

To determine whether the students in the study differed in any statistically significant characteristic other than being taught AD, we tested how well we could predict students' gains from pre-test to post-test based on the variables we had collected. We performed a regression for Gain using Pretest, Instructor, Gender, Final Grade, College, Math, Verbal, and Writing as regressors. The results indicate that none of the variables besides Pretest and Instructor was a factor in a student's gain. Thus, we are confident that the students in the treatment group were not different in any important aspect from the students in the control group.

Comparison of Students by AD Instruction

Our hypothesis was that the students in the first-year writing course who received training in Argument Diagramming would gain significantly more in each category on the two tasks than students who did not receive the training. Since the use of argument diagrams was explicitly taught only by Instructors 2, 6 & 7, this hypothesis was tested by determining whether the average gain of the students taught by Lecturers 2, 6 & 7 was significantly different from the average gain of the students taught by Lecturers 1, 3, 4 & 5. The students taught by Lecturers 2, 6 & 7 are represented in all the tables below by (AD), and the students taught by Lecturers 1, 3, 4 & 5 are represented by (No AD). The mean gain for the sub-populations of students in each treatment group is represented given in Figure 2 for Task 1, and in Figure 3 for Task 2.

To determine the predictive value of AD treatment on a student's gain from pretest to posttest, an ANCOVA was conducted for the gains in each category for Task 1 with AD as a factor and the corresponding pretest score as a covariate. So, for example, we conducted an ANCOVA on the Argument Conclusion Gain with AD as a factor and the Argument Conclusion Pretest as a covariate. The results for Task 1 are given in Table 1.

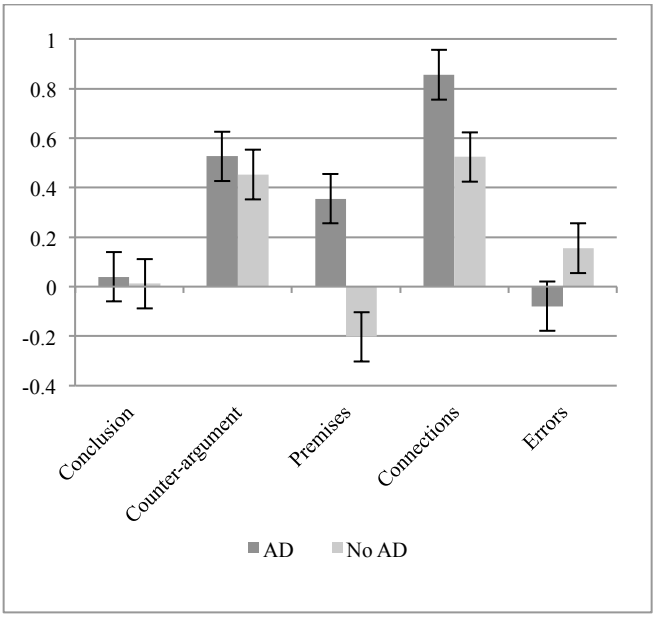

Figure 2: Comparisons of gains in each category of Task 1 from pretest to posttest for students who were and were not taught argument diagramming.

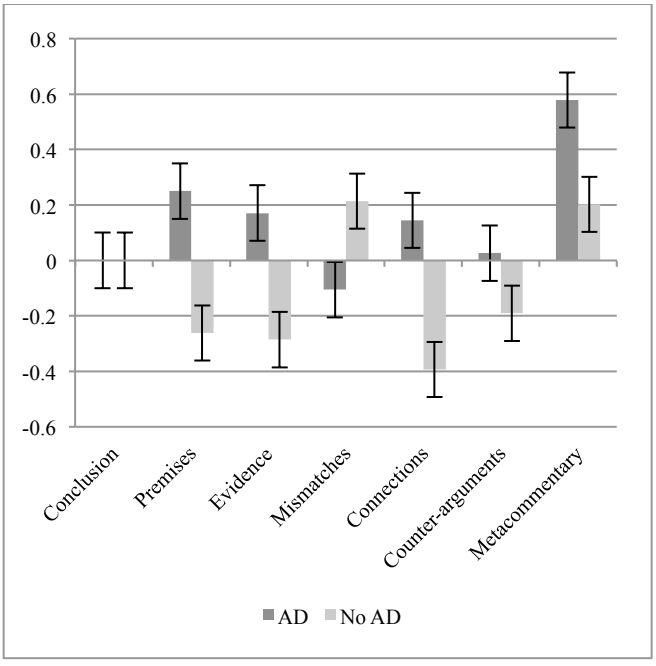

Figure 3: Comparisons of gains in each category of Task 2 from pretest to posttest for students who were and were not taught argument diagramming.

Table 1: ANCOVA test results for the variable AD for each category on Task 1.

Category	$F(1,80)$	p
Argument Conclusion	2.47	0.120
Counter-arguments	0.94	0.335
Premises	4.54	0.036
Connections	7.35	0.008
Errors	6.91	0.010

The effect of AD was statistically significant in each category except Argument Conclusion and Counter-arguments for Task 1.

An ANCOVA was also conducted for the gains in each category for Task 2 with AD as a factor and the corresponding pretest score as a covariate. The results for Task 2 are given in Table 2.

Table 2: ANCOVA test results for the variable AD for each category on Task 2.

Category	$F(1,80)$	p
Argument Conclusion	1.80	0.184
Premises	5.63	0.020
Evidence	6.70	0.012
Mismatches	12.36	0.001
Connections	12.35	0.001
Counter-arguments	5.73	0.019
Metacommentary	10.60	0.002

The effect of AD was statistically significant in each category except Argument Conclusion for Task 2.

Discussion

Findings

The results from Task 1 show that, when reading an argument, students who were taught argument diagramming were significantly more likely than those who were not to identify more of the relevant premises offered that support the author's conclusion, and explain more explicitly how the premises are supposed to work together to support the conclusion. In addition, these students were much less likely to make any errors in their analysis.

The results from Task 2 show that, when evaluating the argument in a text, students who were taught argument diagramming improved significantly more than those who were not in their ability to (a) provide more premises to support their own thesis, (b) offer more evidence in support of each premise (c) have fewer mismatches between premises and evidence, (d) explain more explicitly how the premises are supposed to work together to support the conclusion, (e) offer possible counter-arguments, and (f) provide metacommentary on their response.

Thus, it seems that students who were taught argument diagramming are developing new schema for reading arguments, and learning how to effectively translate this into their own writing. This is reflected most noticeably in the improvement of the metacommentary from pretest to posttest. We conclude that incorporating argument diagramming into the curriculum of *Interpretation and Argument* is positively beneficial to realizing several of our course objectives.

Educational Importance

The primary educational importance of this study is two-fold. First, the results indicate that it is possible to significantly improve students' argumentative writing skills over the course of just one semester, even when the course is not only a critical thinking course. Second, these results indicate that a relatively small addition to the curriculum of a first-year writing course can have dramatic benefits for students. The initial instruction in understanding arguments and creating argument diagrams can be given in one or two class-periods (or an online tutorial) and regular, weekly homework assignments can be added to reading, summary and/or reflection assignments. Supplementing one's teaching with argument diagramming does not require a radical reworking of the syllabus, course readings or assignments. This is a great benefit to instructors who may be reluctant to change a curriculum that has been successful.

Future Work

This study raises as many questions as it answers. While it is clear that the introduction of argument diagramming to the First-Year Writing Program curriculum significantly improves a student's ability to reach several stated course objectives, it would be interesting to explore further the cognitive basis for the effect of argument diagramming. In particular we would like to know what aspects of constructing diagrams help the most in developing new schema.

It would also be interesting to explore whether, once a student learns how to construct argument diagrams, the actual construction of a diagram is important for a particular analysis task. That is, for example, it could be that the new schema is in place, and so the diagrams are no longer needed, or it could be that the construction of a diagram while reading activates the new schema.

We would also like to consider whether there are other skills that we did not measure this time that this addition may help to improve. For example, because our work here did not distinguish between first and second language learners, we cannot speak to whether argument diagramming has more or less of an effect upon second language learners. Additionally, we have anecdotal evidence from several teachers that using argument diagramming during the peer review process was helpful. It would be extremely useful to know whether using argument diagramming in peer review of papers in general makes subsequent drafts better.

Lastly, unlike the relatively solitary activities in which students engage in our FW Courses—like doing homework and writing essays—there are many venues in and out of the classroom in which students may engage in the analysis and evaluation of arguments in a group setting. These may

include anything from classroom discussion of a particular author or topic, to group deliberations about for whom to vote or what public policy to implement. In any of these situations it seems as though it would be advantageous for all members of the group to be able to visually represent the structure of the arguments being considered. We would like to know whether knowing how to construct argument diagrams would aid groups in these situations.

Acknowledgments

I would like to thank three anonymous reviews for their helpful comments. This study was funded by a joint grant from the Spencer Foundation and the Teagle Foundation, under the title "Systematic Improvement of Undergraduate Education in Research Universities."

References

Armbruster, B.B., & Anderson, T.H. (1982). Mapping: Representing informative text graphically. In C.D. Holley & D.F. Dansereau, (Eds.), *Spatial learning strategies*. New York: Academic Press.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.

Bartlett, F. C. (1958). *Thinking: An experimental and social study*. New York: Basic Books.

Dansereau, D.F., Collins, K.W., McDonald, B.A., Holley, C.D., Garland, J., Diekhoff, G., & Evans, S.H. (1979). Development and evaluation of a learning strategy program. *Journal of Educational Psychology, 71*, 64-73.

Enkvist, N. E. (1990). Seven Problems in the Study of Coherence and Interpretability. In U. Connor & A. M. Johns (Eds.), *Coherence in writing: Research and pedagogical perspectives*. Alexandria, VA: Teachers of English to Speakers of Other Languages, Inc.

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice*. New York: W. H. Freeman & Co.

Evans, S. H. (1967). A brief statement of schema theory. *Psychonomic Science, 8*, 87-88.

Fisher, A., & Scriven, D. (1997). *Critical thinking: Its definition and assessment*. CA: Edgepress.

Harrell, M. (2012). Assessing the Efficacy of Argument Diagramming to Teach Critical Thinking Skills in Introduction to Philosophy. *Inquiry, 27*(2), 31-38.

Harrell, M. (2011). Argument Diagramming and Critical Thinking in Introductory Philosophy. *Higher Education Research & Development, 30*(3), 371-385.

Harrell, M. (2008). No Computer Program Required: Even Pencil-and-Paper Argument Mapping Improves Critical Thinking Skills. *Teaching Philosophy, 31*, 351-374.

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*(1), 95–111.

Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.

Johnson, D.D., Pittleman, S.D., & Heimlich, J.E. (1986). Semantic mapping. *Reading Teacher*, 39, 778-783.

Kirschner, P.A., Shum, S.J.B., & Carr, C.S. (Eds.). (2003). *Visualizing argumentation: Software tools for collaborative and educational sense-making*. New York: Springer.

Kuhn, D. (1991). *The skills of argument*. Cambridge: Cambridge University Press.

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

Mandler, J. M. (1984). *Stories, scripts, and scenes: Aspects of schema theory*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Moore, D. W., & Readence, J. E. (1984). A quantitative and qualitative review of graphic organizer research. *Journal of Educational Research, 78*(1), 11-17.

Novak, J.D., & Gowin, D.B. (1984). *Learning how to learn*. New York: Cambridge University Press.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Shapiro, B. C. Bruce & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education* (pp. 33-58). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Scheuer, O., McLaren, B., Harrell, M., & Weinberger, A. (2011). Will Structuring the Collaboration of Students Improve Their Argumentation? In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Eds.), *Lecture notes in computer science: Artificial intelligence in education—15th international conference.* Berlin: Springer-Verlag.

Schwartz, R.M., & Raphael, T.E. (1985). Concept of definition: A key to improving students' vocabulary. *Reading Teacher*, 39: 198-205.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295-312.

Twardy, C.R. (2004). Argument maps improve critical thinking. *Teaching Philosophy, 27*, 95-116.

van Gelder, T., Bissett, M., & Cumming, G. (2004). Cultivating expertise in informal reasoning. *Canadian Journal of Experimental Psychology, 58*, 142-152.

Vande Kopple, W. (1989). *Clear and coherent prose: A functional approach.* Glenview, IL: Scott, Foresman, & Company.

Winn, W. (1991). Learning from maps and diagrams. *Educational Psychology Review*, 3: 211-247.