A Meta-Analysis of Graphic Organizer Reading Interventions for Learning Science and Social Studies Content

Ashley Sandoval
amoorshead@smu.edu

Follow this and additional works at: https://scholar.smu.edu/simmons_dtl_etds

Recommended Citation
Sandoval, Ashley, "A Meta-Analysis of Graphic Organizer Reading Interventions for Learning Science and Social Studies Content" (2020). Department of Teaching and Learning Theses and Dissertations. 7. https://scholar.smu.edu/simmons_dtl_etds/7

This Dissertation is brought to you for free and open access by the Department of Teaching and Learning at SMU Scholar. It has been accepted for inclusion in Department of Teaching and Learning Theses and Dissertations by an authorized administrator of SMU Scholar. For more information, please visit http://digitalrepository.smu.edu.
DISSENTATION APPROVAL

This dissertation submitted by _____ has been read and approved by the following faculty members of the Annette Caldwell Simmons School of Education and Human Development at Southern Methodist University. The final copy has been examined by the Dissertation Committee and the signatures which appear here verify the fact that any necessary changes have been incorporated and that the dissertation is now given the final approval with reference to content, form and mechanical accuracy.

The dissertation is therefore accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Advisor and Committee Chair: [Stephanie Al Otaiba, Ph.D.]
Signature: ___________________________ Date: ________________

Co-Chair: [Amy Gillespie Rouse, Ph.D.]
Signature: ___________________________ Date: ________________

Committee Member: [Paul Yovanoff, Ph.D.]
Signature: ___________________________ Date: ________________

Committee Member: [Steve Ciullo, Ph.D; Texas State University]
Signature: ___________________________ Date: ________________
A META-ANALYSIS OF GRAPHIC ORGANIZER READING INTERVENTIONS FOR LEARNING SCIENCE AND SOCIAL STUDIES CONTENT

A Dissertation Presented to the Graduate Faculty of

the Simmons School of Education and Human Development

Southern Methodist University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Education by

Ashley Moorshead Sandoval

July 29, 2020
ABSTRACT

In this meta-analysis, the effects of text-based graphic organizer interventions on content learning in Science and Social Studies for students in grades 3-12, including typically developing and struggling students, and students with disabilities (i.e., LD, ID, and ASD) were assessed. A random-effects model with a correlated robust variance estimation approach was used and showed text based graphic organizer interventions had a large overall positive effect (ES = 1.65) on content learning in both Science and Social Studies. In addition, moderator analysis indicated that text-based graphic organizer interventions were more effective when students received teacher instruction prior to completing graphic organizers independently. Thus, an important implication of this meta-analysis is that text-based graphic organizers are an effective intervention for content learning for a wide range of students, particularly when students are provided with teacher instruction on how to complete the graphic organizer prior to using them independently. Limitations and implications for future research are provided.
ACKNOWLEDGEMENTS

I would like to acknowledge the many people who helped me complete this dissertation. I am particularly grateful to Stephanie Al Otaiba, my advisor and dissertation committee co-chair, for her direction and encouragement. Thank you for showing me what excellence in research and education is. I have learned so much about research and writing because of the time you have spent with me providing resources, ideas, and words of encouragement. I would also like to thank Amy Gillespie Rouse for being my dissertation co-chair and for the hours you have spent guiding me as my meta-analysis mentor. Thank you to my other dissertation committee members, Steve Ciullo and Paul Yovanoff for your support and expertise. I am also grateful to all the faculty I have had the pleasure of working with and learning from while at SMU, particularly Jill Allor, for giving me the opportunity to work on Project Intensity. I would also like to thank my parents for teaching me the value of education and investing wholeheartedly in me and my education. Finally, I would like to thank my husband, Barry, for being my unwavering support system and for being the best editor I could ask for.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>I. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Text Complexity in Science and Social Studies</td>
<td>2</td>
</tr>
<tr>
<td>Why Students with Disabilities May Struggle with Reading to Learn</td>
<td>3</td>
</tr>
<tr>
<td>Cognitive Models of Learning</td>
<td>4</td>
</tr>
<tr>
<td>Cognitive Models of Reading Comprehension</td>
<td>9</td>
</tr>
<tr>
<td>A Cognitive Model of Reading to Learn</td>
<td>10</td>
</tr>
<tr>
<td>Cognitive Processing Issues when Reading to Learn</td>
<td>12</td>
</tr>
<tr>
<td>General and Disciplinary Literacy Strategy Instruction</td>
<td>14</td>
</tr>
<tr>
<td>Graphic Organizers and Reading to Learn</td>
<td>16</td>
</tr>
<tr>
<td>Previous Graphic Organizer Research</td>
<td>20</td>
</tr>
<tr>
<td>Purposes for this Meta-Analysis</td>
<td>22</td>
</tr>
<tr>
<td>Research Questions</td>
<td>25</td>
</tr>
<tr>
<td>II. Method</td>
<td>27</td>
</tr>
<tr>
<td>Study Inclusion Criteria</td>
<td>27</td>
</tr>
<tr>
<td>Search Strategies</td>
<td>27</td>
</tr>
<tr>
<td>Study Coding</td>
<td>31</td>
</tr>
<tr>
<td>Inter-Rater Reliability</td>
<td>32</td>
</tr>
<tr>
<td>Statistical Analysis</td>
<td>32</td>
</tr>
<tr>
<td>Descriptive Analysis of Quality Indicators</td>
<td>36</td>
</tr>
<tr>
<td>III. Results</td>
<td>37</td>
</tr>
<tr>
<td>Description of Studies</td>
<td>37</td>
</tr>
<tr>
<td>Descriptive Analysis of Quality Indicators</td>
<td>41</td>
</tr>
<tr>
<td>Research Question One</td>
<td>42</td>
</tr>
<tr>
<td>Research Question Two</td>
<td>43</td>
</tr>
<tr>
<td>Publication Bias</td>
<td>43</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Study Characteristics, Study Quality, and Average Effect Sizes</td>
<td>61</td>
</tr>
<tr>
<td>2. Student and Intervention Characteristics</td>
<td>63</td>
</tr>
<tr>
<td>3. Meta-Regression Moderator Analysis</td>
<td>67</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. PRISMA Chart of Study Search and Article Eligibility/Selection Procedures</td>
<td>68</td>
</tr>
<tr>
<td>2. Forest Plot of the Average Effect Size of Individual Studies</td>
<td>69</td>
</tr>
<tr>
<td>3. Funnel Plot for Publication Bias</td>
<td>70</td>
</tr>
</tbody>
</table>
CHAPTER ONE
INTRODUCTION

Science and Social Studies literacy are integral parts of academic achievement and active
citizenship (Ciullo et al., 2020; Cromley, 2009). Yet student performance on national measures
of content knowledge in Science and Social Studies provides insight into the degree to which
students struggle with content learning. In 2015, the U.S. Department of Education reported that
on national assessments of Science knowledge, only 38% of fourth-grade students assessed at or
above the proficient level in Science, 34% of eighth-grade students assessed at or above
proficient, and 22% of students in twelfth grade demonstrated proficiency (NAEP, 2015). Even
within these low percentages, students without identified disabilities had higher levels of
achievement than students with disabilities. Only 18% of fourth graders, 11% of eighth graders,
and 9% of twelfth graders at the proficient level were students with identified disabilities
(NAEP, 2015).

Similarly, students across the country have demonstrated low levels of performance on
national measures of Social Studies content knowledge. On the 2014 NAEP U.S. History
Assessment, only 18% of eighth-grade students performed at or above the proficient level in U.S.
History, 27% were at or above the proficient level in Geography, and 23% of students were at or
above a proficient level in Civics (NAEP, 2014). Students with disabilities had lower levels of
proficiency than non-disabled peers, as only 5%, 8%, and 8% of students with disabilities scored
at or above the proficiency level on U.S. History, Geography, and Civics assessments,
respectively (NAEP, 2014).
The pervasive difficulties students demonstrate with learning content, particularly for students with disabilities, indicate that intervention is necessary. Given that reading is the primary vehicle for content learning in grades 4-12 (Gajria et al., 2007; Swanson et al., 2014), to implement effective interventions it is important to view reading and learning holistically and consider the relationship between three factors: the complexity of the text, the ability of the reader, and the task of reading to learn, as when these components are aligned, comprehension and ultimately learning occur (Snow, 2002; Snow, 2018). If the complexity of the text is too great, the task is not understood, or the ability of the reader is too weak (or any combination of these factors is not met), comprehension fails, and ultimately learning also fails (Snow, 2018).

**Text Complexity in Science and Social Studies**

Science texts are complex, and Science achievement has been positively correlated with reading ability (Cromley, 2009). Science texts are often challenging to comprehend given their complex structure, unique genre, detailed explanations, use of argument, and inclusion of multiple documents and sources (Britt et al., 2014). In addition, the conceptual density and use of unfamiliar technical vocabulary in Science texts also make comprehending and learning from reading challenging (Mason & Hedin, 2011).

It is widely recognized that textbook reading is the most commonly used tool for Social Studies learning (Okolo & Ferretti, 1996). Yet, Social Studies textbooks are particularly challenging to read given their complexity. Specifically, the following qualities of Social Studies texts add to their overall text complexity: they lack conceptual coherence, they tend to cover a breadth not depth of subjects, and they often contain irrelevant or confusing images and visual displays of information (Okolo & Ferretti, 1996). Social Studies texts also present comprehension and learning challenges for students, as they are often written at readability levels
that are beyond students’ independent reading levels, and these texts are also often written with poorly organized text structures and information (Swanson et al., 2014).

Thus, given the prevalence and importance of learning from textbooks in both Science and Social Studies, enabling students to understand and learn from complex texts is not only a reading issue, but it also becomes an issue with having appropriate opportunities to learn (Cawley & Parmar, 2001). Students cannot be held accountable for achievement in the disciplines until educators have ensured that students have been given adequate and appropriate opportunities to become skillful in their abilities to read to learn. Therefore, in order to provide students with adequate and appropriate opportunities to learn from complicated Science and Social Studies texts, educators must not only recognize the complexity of Science and Social studies text, but they must also identify which students struggle when reading to learn, as well as deliver effective interventions to enable all students the opportunity to learn while reading.

**Why Students with Disabilities May Struggle with Reading to Learn**

While increased demands from content area texts provide challenges for all students, content area texts provide an even greater challenge for those students who demonstrate difficulties with comprehension and identified disabilities. This includes students with identified learning and intellectual disabilities, as well as students who been identified with Autism Spectrum Disorder.

Low literacy levels interfere with learning in content area subjects, and poor readers struggle to learn in reading intensive courses (Biancarosa & Snow, 2006). Struggling readers have difficulties attaining and retaining basic concepts, as well as developing background knowledge while reading informational text, which may interfere with their ability to create and analyze complex ideas (Cain et al., 2001; Ciullo et al., 2016).
Learning from content area texts may be particularly problematic for students with learning disabilities (LD) as they often read below grade level and struggle with expository text structure and also struggle with academic and content vocabulary (Barton-Arwood & Little, 2013). For these students, comprehension problems are complex and often a combination of one or more components of reading. Comprehension difficulties for these students may be the result of one or more of the following factors: low vocabulary knowledge, weak fluency, inadequate conceptual knowledge, inability to be strategic when comprehending, or difficulty with reasoning or making inferences (Roberts et al., 2008).

Learning from text is also challenging for students with Autism Spectrum Disorder (ASD). Students with ASD demonstrate comprehension difficulties, specifically organizing content, connecting ideas, recognizing text cues, and self-monitoring when comprehending (El Zein et al., 2013). Since students with ASD consistently demonstrate appropriate word recognition skills, but maintain poor reading comprehension, it is hypothesized that language deficits, not decoding deficits, interfere with appropriate comprehension in students with ASD (El Zein et al., 2013).

There appears to be no published research on children with intellectual disabilities (ID) learning from text. Research on interventions to teach general reading comprehension is limited, and comprehensive studies that involve multi-component reading interventions are lacking (Browder et al., 2006). Thus, further research on effective reading comprehension interventions for students with ID is needed.

**Cognitive Models of Learning**

Cognitive psychology is one perspective that can be used to create a unified understanding of reading and learning, or reading to learn, to help develop identify and develop
appropriate interventions for struggling readers (Dole & Nokes, 2009; McMaster et al., 2014).

Within the cognitive perspective, reading and learning are both viewed as series of component mental processes, or cognitive strategies, that individuals use to process information (Dole & Nokes, 2009). Learning and reading comprehension are achieved when a coherent mental representation of text or learned content is created (McMaster et al., 2014). Within the general cognitive perspective of learning and reading, the Expertise view of learning, Van den Broek et al.’s (2017) cognitive theory of reading comprehension, and Schema Theory can be combined to create a comprehensive cognitive theory of reading to learn.

**Expertise Theory, Learning, and Knowledge Representation**

In contemporary theories of disciplinary learning, the ultimate goal of learning is the development of expertise. Experts differ from others, as they hold more knowledge about their given domains, and the knowledge that they hold is highly structured (Chi, 2006). As a result of holding more sophisticated and more highly organized structures of knowledge, experts are able to reason and solve problems more efficiently and effectively (Chi, 2006). On the other end of the expertise continuum is the novice who is inexperienced within a domain and holds a fragmented representation of knowledge and who, as a result, is unable to solve problems effectively and efficiently (Chi, 2006).

The majority of students can develop expertise in a domain through learning, studying, and deliberate practice (Chi, 2006). Learning, or the development of expertise, can be understood across a developmental continuum (Chi, 2006). Students move from being a novice (someone with minimal experience in a domain) to an initiate (someone who has begun introductory instruction; usually beginning in K-12). Next they move from being an apprentice (a student participating in instruction beyond the introductory level; often occurring in post-
secondary instruction), to a journeyman (a high level of competence, demonstrated by experience and reliable work), and finally to the position of expert. Each stage along the continuum of expertise represents a mental structural change, with individuals demonstrating increasingly more sophisticated and more highly organized representations of knowledge at each level (Chi, 2006). Students enter into primary school as naïve, or as one who is totally ignorant of a domain, or as one who may hold naïve beliefs or misconceptions about a domain (Chi, 2006). As students move to higher grade levels and learn more content, they ideally develop more sophisticated ways of representing knowledge that will then position them to move into the next level of the expertise continuum.

Types of Knowledge. Cognitive psychologists posit that there are two dominant types of complex knowledge—declarative knowledge and procedural knowledge. Procedural knowledge is knowing how and is considered prescriptive and use-specific (Chi & Ohlsson, 2005). It is believed procedural knowledge is represented as scripts, or schemas for events (Byrnes, 2001). In contrast to procedural knowledge, declarative knowledge is the knowledge of concepts, principles, ideas, and schemas and is often described as “knowing what” (Chi & Ohlsson, 2005).

Experts hold sophisticated models of declarative and procedural knowledge, however when students are reading to learn, acquisition of declarative or conceptual knowledge is the focus. Specifically, declarative knowledge, or conceptual knowledge, is the type of knowledge typically referenced when describing the types of knowledge students are expected to acquire while reading expository texts and textbooks. An example of declarative knowledge in Social Studies would be understanding important concepts and events, as well as the relationships between these events, whereas an example of declarative knowledge in Science would be
understanding concepts such as motion and force and understanding the unifying theory that explains the relationship between them (Chi, 2006).

**Conceptual Knowledge and Expertise.** A foundational assumption of the Expertise View of cognition is that knowledge is represented qualitatively differently between experts and novices, particularly in formal domains. First, experts hold a larger body of conceptual knowledge than novices. Second, expert knowledge is organized in a way that is more sophisticated than the way novices structure information (Chi, 2006). That is, experts see patterns and features in information that are not apparent to novices, and knowledge is often represented around key theories of a domain (Chi, 2006). For example, Chi et al. (1981) have shown that expert physicists represent knowledge around core principles, whereas novices represent information around independent situations or individual formulas. Differences in performance between experts and novices is determined by the way in which knowledge is represented (Chi, 2006).

**Representations of Conceptual Knowledge.** There are three dominant theories or ways conceptual domain knowledge is believed to be represented: semantic networks, theories, and schemas (Chi, 2006; Chi & Ohlsson, 2005; Markman, 1999). The semantic network view of knowledge representation hypothesizes that knowledge is organized in semantic networks. These semantic networks are represented by links and nodes. Nodes are the mental representation of concepts, and links define the relationships between these concepts. Storing information requires learners to create inferences about the relationships between information, and information is retrieved by scanning these inferences/relationships, or links, between concepts. From the Expertise view of cognition, since there is an infinite number of nodes and links organizing one’s declarative knowledge, it is assumed that processing efficiency is
increased as these concepts and relationships are grouped by domain (Chi & Ohlsson, 2005). Domains can be informal, such as ordering food from a restaurant, or formal, such as Biology knowledge (Chi & Ohlsson, 2005).

It is theorized that formal domain knowledge is mentally represented in two ways (Chi & Ohlsson, 2005). First, information can be represented around dominant relationships. These relationships include: a hierarchical relationship, a cause-effect relationship, a process or before/after representation, or a locally structured relationship. Domains of declarative knowledge, particularly in formal domains, can also be organized around key theories. As Chi and Ohlsson (2005) explain, “a theory is organized around a small set of core concepts or principles—big ideas—on which the rest of the elements in the domain are dependent” (p. 6). Examples of theories in Science include the theory of relativity or Darwin’s theory of evolution (Chi & Ohlsson, 2005). For Social Studies, an example of a big idea or guiding principle is the idea of continuity and change or cause and consequence (Historical Thinking Project, n.d).

**Schema Theory**

Schema Theory is a related theory of how conceptual information is structured in cognition. This theory is based on the assumption that conceptual knowledge is represented by the reoccurring patterns and instances from personal experience. Chi and Olson describe a schema as, “a set of relations that amount to a set of slots or attributes, where the slots can be thought of as variables that can take values within a specified range.” (p. 6). Schemas are bundles of interwoven information, and it is believed that they are retrieved as these bundles (Chi & Ohlsson, 2005).

It is believed experts and novices differ with regard to their schemas, with experts chunking domain knowledge around meaningful patterns. For experts, domain knowledge is
not haphazardly stored as individual facts and formulas. Rather, knowledge is coherently organized around core concepts, meaningful relationships, and consistent mental models (Bransford et al., 2000).

**Cognitive Models of Reading Comprehension**

Similar to cognitive learning theorists, cognitive reading theorists also hold that knowledge representation is the foundation of effective processing and comprehension of text. Van den Broek et al. (2017) note that while the specific focus of cognitive theories of comprehension hold differing views about how the component cognitive processes of comprehension interact, there is general consensus that the cognitive processes of knowledge representation in reading comprehension can be categorized into two levels. The most basic level of knowledge representation in reading involves representation of words, syntax, and relationships between sentences, and the second higher-order representation of information text is represented more globally (van den Broek et al., 2017). At the higher level, text is represented as elements, such as facts, events, or individuals, and is connected by semantic relationships: causal, referential, logical, spatial, temporal, associative (van den Broek et al., 2017). In addition, this higher level of representation requires selecting information through attention allocation, executive functions, and inference-making. It is during this higher-order process that concepts in text are processed into meaningful and cohesive mental representations, and understanding of text occurs (Kendeou et al., 2014).

**Schema Theory and Knowledge Representation in Reading**

Schema Theory is another cognitive view of reading used to explain how information is represented at the highest level of comprehension. While the basic principles of Schema Theory overlap with cognitive views of schema previously described, reading researchers who support
Schema Theory differ in that the dominant model of text representation, or text schema, involves the type of text (i.e., narrative or expository) being read. Meyer (1985), one of the most widely cited schema theorists, posits that readers create schemas around the prose structures, or the dominant organizational relationships between concepts in a text. Common expository prose structures include: classification, causation, response, comparison, and description (Byrnes, 2001).

A Cognitive Model of Reading to Learn

Expertise Theory, Van den Broek et al.’s (2017) cognitive model of reading, and Schema Theory can be combined to create a unified cognitive model of reading to learn. While cognitive models of learning and reading comprehension have to a large extent been developed separately, it is widely agreed from both perspectives that representation of domains of organization is organized in one of two ways. The first is that information is believed to be cognitively structured around dominant or salient relationships (e.g., cause and effect, classification, compare and contrast). This view is held by reading schema theorists and general cognitive theorists. The second view, held by Expertise cognitive theorists, is that domains of knowledge are organized around key theories or principles, such as continuity and change, cause and consequence in Social Studies, or laws and theories in Science. Thus, the instruction goal of interventions for reading to learn should be to enable students to create organized mental models of learned knowledge.

While organized knowledge is the product of reading to learn, it is also important to have a cohesive model of the cognitive processes involved in creating mental representations. Borrowing from Van den Broek et al. (2017), when students are reading to learn they are first representing written text into meaningful language units as they decode words, engage in
fluency, and understand vocabulary. If a student is proficient with these lower level processes his or her working memory will not be overloaded and the student will be able to execute higher-order comprehension processes, or reading to learn (Kendeou et al., 2014).

Cognitive perspectives of higher-order processing in reading comprehension and learning overlap, as both frameworks posit that the higher levels of comprehension and learning processes both involve three distinct sub-processes: selecting information, organizing information, and integrating information, all of which impact processing efficiency of this information (Mayer, 1997; van den Broek et al., 2017).

Selection of information while reading to learn is defined as the act of focusing attention to key information (McCrudden & Rapp, 2017). As readers become more proficient, they develop greater ability to focus on the key components of text and ignore less important information (van den Broek et al., 2017). Van den Broek et al. (2017) call this ability to be more selective and efficient in attending to information sensitivity to structural centrality. As a result, the selected information gains a more prominent position in the mental representation of the text.

Organization is the next level of cognitive processing and is defined as creating inferences between and within information (McCrudden & Rapp, 2017). Van den Broek et al. (2017) note that inferences enable readers to connect information in a meaningful way. More sophisticated readers are able to draw inferences from larger chunks of text (i.e., paragraphs and sections versus individual ideas expressed in sentences) and they have a more sophisticated ability to create abstract inferences around themes and unobservable traits such as beliefs and feelings (van den Broek et al., 2017). Inferences are the act of creating semantic connections or links between nodes of knowledge, described previously.
Integration acts as a complement to organization and is defined as the process of assimilating key information and prior knowledge (McCrudden & Rapp, 2017). It is postulated that integration can be active or passive. Active integration involves automatic integration, whereas active integration is guided by the learner (McCrudden & Rapp, 2017).

Cognitive Processing Issues When Reading to Learn

Students who struggle with reading to learn may have processing difficulties at the higher level of processing, the lower level, or both. Students who struggle with higher levels of processing may be unable to identify the main idea of a text, organize key details, or make a coherent representation of information (Kendeou et al., 2014). From a cognitive viewpoint, it is assumed that students who display these reading difficulties are experiencing challenges creating coherent mental representations of text as a result of being unable to make inferences, allocate attention, or assimilate key knowledge or background knowledge (Kendeou et al., 2014). Students who have comprehension difficulties at the lower level will demonstrate issues with phonological process and decoding and will be challenged when mentally coding written units of language meaningfully. As a result of working memory limitations and incomplete representations of written code, these students will also be unable to engage in higher and more sophisticated processing of text (Kendeou et al., 2014).

Issues with Higher-Level Processing While Reading to Learn

As children become more proficient in reading, they become more sophisticated and efficient in their ability to attend to the central message of text (Kendeou et al., 2014). They are able to select relevant information even when the text increases in length or becomes denser. However, children who struggle with selection or attention-allocation may be more distracted or overwhelmed by details and unable to focus on the central message of a text (Kendeou et al.,
2014); they may experience difficulty comprehending text with breaks in coherence, which leads to less structured and cohesive mental representations (Kendeou et al., 2014).

Students who struggle with executive functioning struggle with higher level comprehension and integration of information, as a result of working memory and inhibition issues (Kendeou et al., 2014). Inhibition allows for the suppression of irrelevant information to allow for selection and attention to relevant information such as key details and the central message of a text. When inhibition does not occur, working memory becomes overloaded and information is not integrated.

Working memory capacity increases as individuals age, which results in increased inference making and comprehension monitoring (Kendeou et al., 2014). Students who struggle with integration and memory-capacity are limited in how much information can be held in working memory, which negatively impacts inference generation and comprehension monitoring (Kendeou et al., 2014). In addition, limitations with integration and working-memory negatively impact the cohesiveness and the quality of the mental representation made of learned or read material.

Inference making is essential to connecting ideas when representing text while reading to learn. As children age, they become better able to make inferences across larger chunks of text and are able to infer more abstract relationships (Kendeou et al., 2014). Students who struggle with making inferences are unable to infer relationships between ideas and are unable to connect ideas within sentences, between sentences and paragraphs, or both (Kendeou et al., 2014). In addition, if a reader is unable to understand the goal text, insufficient and inadequate inferences can be made (Kendeou et al., 2014). Lastly, inference making can be insufficient and inadequate if a student lacks necessary background knowledge, such as concept/content knowledge, or if the
student lacks necessary background knowledge about how information is structured and organized (e.g., text structure) (Kendeou et al., 2014). Insufficient or incomplete representations of knowledge occur when students are unable to make appropriate inferences.

**General and Discipline Literacy Strategy Instruction**

In the past decade, two approaches to reading to learn have emerged: general literacy and disciplinary literacy. In both Science and Social Studies, general literacy content area-instruction focuses on strategies such as learning key vocabulary, comprehension strategy instruction, text structure instruction, strategic questioning, and peer-mediated learning (Ciullo et al., 2020; Reed et al., 2016). General literacy theorists believe the advantage of teaching broadly applicable strategies is that they can be easily transferred across disciplines, and it is believed that because of the universality of these strategies, they are most beneficial for struggling readers (Reed et al., 2016).

In contrast to the content general approach, the disciplinary literacy approach is loosely based on Expertise learning theory, outlined previously. The goal of disciplinary literacy instruction is to help students approach text in a parallel manner to disciplinary experts and focuses on the unique aspects of disciplinary text (Shanahan & Shanahan, 2008). In Science, discipline-specific aspects of reading, learning, and understanding scientific arguments have been identified as: analyzing and integrating visual representations of information, establishing causal relationships between ideas, using scientific language and vocabulary, comprehending scientific arguments, joining claims to evidence, and explaining scientific phenomena specifically and precisely (Reed et al., 2016).

The disciplinary approach to History prioritizes critical thinking and the development of historical thinking (MacArthur et al., 2002). Historical thinking emphasizes historicism versus
presentism (i.e., analyzing past actions, events, and behavior through culture, opportunity, and environment). Through this perspective, historical argument is also emphasized and requires a knowledge of general argument, as well as discipline-specific knowledge such as consideration of sources (MacArthur et al., 2002). Discipline-specific aspects of reading and learning in Social Studies include source analysis, evaluation of author perspective and evidence, and contextualization (Ciullo et al., 2020).

The place of disciplinary literacy within the developmental literacy trajectory is somewhat muddled. Basic literacy is known to occur prior to general literacy instruction, however the relationship between disciplinary literacy and general literacy is not clear. One assumption of disciplinary literacy instruction is that it should occur after students are proficient in general literacy, which students are expected to be proficient in by the end of middle school (Shanahan & Shanahan, 2008). However, disciplinary literacy is also sometimes presented as a replacement for general literacy, where students are expected to move from basic literacy directly into literacy with a disciplinary focus (Faggella-Luby et al., 2012).

While the theoretical alignment of disciplinary literacy is sound, it is problematic to focus on high level reasoning, argument, and synthesizing information without considering that knowledge representation is an essential foundational building block of this higher level comprehension and learning. This assumption and ultimately approach to reading in the content areas is particularly problematic for students with disabilities, who may struggle with both low and high levels of information processing and consequently struggle with representing information.

Disciplinary Literacy theorists have largely overlooked that knowledge representation is the foundation of reading to learn in their model. They have not considered that students who
cannot represent content knowledge effectively cannot engage in higher order thinking involved in reasoning, argument, and synthesis. They have failed to create a model that enables students to build a strong foundational knowledge of representing information while reading to learn and they have also failed to create a model that supports students who struggle mentally representing and processing content text. Thus, before students can engage effectively in disciplinary literacy, they need to be provided with strategies that enable them to learn content material and need to be taught how to select and organize information using core principles and themes of a discipline as they read to learn.

**Graphic Organizers and Reading to Learn**

Graphic organizers are one way to develop content knowledge through knowledge representation. They have shown improved outcomes for students, particularly students with LD (Dexter & Hughes, 2011; Gajria et al., 2007). Graphic organizers are defined as visual representations of information that highlight the key structural relationships between core concepts and help students create mental representations of knowledge. (Ciullo et al., 2014; Darch & Eaves, 1986; Dexter & Hughes, 2011; Gajria et al., 2007; Kim et al., 2004). Graphic organizers offer learners and readers a template or scaffold and make the implicit cognitive relationships used to structure and organize information explicit.

Graphic organizers facilitate the higher levels of cognitive processing in reading comprehension and learning by supporting readers with the processes of inference-making, executive function, and attention-allocation (McCrudden & Rapp, 2017). Specifically, graphic organizers support readers and learners as they select, organize, and integrate information, which leads to greater processing efficiency with regard to attention and working memory (McCrudden & Rapp, 2017).
Graphic organizers help facilitate selection of information through the use of colors, arrows, and shapes which act as signals or cues, highlighting more important information and drawing learners’ attention away from less important information, in turn impacting how this information is processed (McCrudden & Rapp, 2017). The process of directing attention towards important information and consequently highlighting essential information which supports effective processing of information, is called the signaling principle (McCrudden & Rapp, 2017). McCrudden and Rapp (2017) further note that while the signaling principle explains the rationale for using visual cues in graphic organizers to highlight important information, signaling independently does not necessarily promote learning. Thus, organization and integration are also essential to learn.

Graphic organizers also facilitate organization or the creation of mental representation of information by directing students to the types of organizational inferences that should be made between information (McCrudden & Rapp, 2017). For example, semantic maps and taxonomies encourage hierarchical inferences and promote information to be coded and organized by superordinate and subordinate relationships (Dexter & Hughes, 2011; McCrudden & Rapp, 2017). Semantic feature analysis, diagrams, and matrices promote the coding of relational inferences and help students organize information around comparative relationships (Dexter & Hughes, 2011; McCrudden & Rapp, 2017). Finally, timelines and flowcharts highlight the chronological ordering of information and the temporal relations between concepts and connect information through temporal inferences (Dexter & Hughes, 2011; McCrudden & Rapp, 2017). These features of graphic organizers can support struggling readers in developing more cohesive mental representations of texts in Science and Social Studies.

*Types of Graphic Organizers*
Contemporary graphic organizers can be divided into three main types, each corresponding with contemporary views of knowledge representation in cognitive science. Appendix B shows examples of each of the types of graphic organizers included in this meta-analysis.

**Concept Maps or Cognitive Maps.** Concept maps are one type of graphic organizer developed by Novak to capture conceptual change of children’s conceptual understanding of Science (Novak, 2010). Concept maps represent domain knowledge, are characterized by concepts and semantic units, and most directly align with the semantic network theory of how information is organized. In concept maps, concepts are visually represented by being enclosed by a box or circle and are connected by a semantic unit (Novak & Cañas, 2006). Furthermore, the relationships in concept maps are arranged in a hierarchical fashion, with the most general or inclusive attributes or concepts at the top and the most specific details or concepts at the bottom (Novak & Cañas, 2006).

Cognitive maps are similar to concept maps in that major ideas are positioned and linked by lines and arrows to show key relationships (Darch & Eaves, 1996; Dexter & Hughes, 2011). In addition, key details are highlighted through the use of reduced language and keywords (Dexter & Hughes, 2011). Cognitive maps differ from concept maps in that prepositions are not used to link key phrases.

**Semantic Maps and Semantic Feature Analysis.** Semantic maps highlight hierarchical relationships, organize information in terms of superordinate and subordinate relationships, and help students code information through hierarchical inferences (Dexter & Hughes, 2011; McCrudden & Rapp, 2017). Semantic feature analysis graphic organizers highlight the
comparative relationships between information and help students code information through relational inferences (Dexter & Hughes, 2011; McCrudden & Rapp, 2017).

**Text Structure or Modal Graphic Organizers.** Meyer’s (1985) theory of schema for expository text and general domain views of information aligns with the dominant types of graphic organizers in contemporary reading comprehension literature. These graphic organizers are often categorized by the type of expository text structure they represent: narration, compare/contrast, process, problem/solution, or cause/effect (Akhondi et al., 2011). Dexter and Hughes (2011) call these modal graphic organizers visual displays, which include: timelines, decision trees, flow charts, taxonomies, and Venn diagrams. Timelines and flowcharts highlight the chronological ordering of information and the temporal relations between concepts and connect information through temporal inferences (Dexter & Hughes, 2011; McCrudden & Rapp, 2017). Taxonomies, while visually different than semantic feature analysis graphic organizers, also organize information in terms of superordinate and subordinate relationships and help students code information through hierarchical inferences (Dexter & Hughes, 2011; McCrudden & Rapp, 2017). Venn diagrams and matrices are other types of modal graphic organizers that highlight the comparative relationships between information and help students code information through relational inferences (Dexter & Hughes, 2011; McCrudden & Rapp, 2017).

It is important to note that while there is increased interest in disciplinary approaches to text in the contemporary literature, there is very little mention of disciplinary-specific graphic organizers in the literature. This further evidences the importance of a unified cognitive model of reading that focuses on teaching students how to represent conceptual knowledge using disciplinary themes and core concepts when reading.
Previous Graphic Organizer Research

**Graphic Organizers and Reading Comprehension**

Kim et al. (2004) synthesized the effects of graphic organizers (defined as semantic organizers, outlines, and cognitive maps with and without mnemonics) on reading comprehension improvements for students with LD. Kim et al. (2004) reported that students who received graphic organizers during intervention had higher comprehension scores than students receiving traditional instruction. Kim et al. (2004) reported large effect sizes across studies, whether a researcher, teacher, or researcher-teacher led the graphic organizer reading comprehension intervention. No significant differences in effect sizes were found when comparing students across grade level when students received a graphic organizer intervention.

**Graphic Organizer Research in Content Area Reviews**

Science and Social Studies graphic organizer research has been embedded in systematic reviews and meta-analyses with content learning. Previous meta-analyses and syntheses have reported moderate to large effect sizes for content learning in Science and Social Studies using graphic organizers that represent knowledge as a general domain or using expository text schema. In their systematic review, Ciullo and Reutebuch (2013) reported high effects of Social Studies learning when computer and text-based graphic organizers were used in interventions for students with LD.

**Content Area Graphic Organizer Meta-Analyses**

The two most-cited disciplinary focused graphic organizer meta-analyses are from Dexter and Hughes (2011) and Dexter et al. (2011). Dexter and Hughes (2011) evaluated the effects of
graphic organizer interventions on content learning, vocabulary knowledge, and comprehension in English/reading, Science, Social Studies, and Mathematics for students with LD in grades 4-12. Dexter et al. (2011) evaluated the effects of graphic organizer Science interventions on vocabulary knowledge and factual comprehension.

Dexter and Hughes (2011) reported a large overall effect of graphic organizers on the content and vocabulary learning and comprehension for students with LD in upper elementary through high school ($ES = .91$). Similarly, Dexter et al. (2011) reported a large overall standardized effect of graphic organizers on Science learning (measured by comprehension and vocabulary assessments) ($ES = 1.052$).

Dexter and Hughes (2011) reported differential effects for type of graphic organizer (i.e., cognitive mapping, semantic mapping, semantic feature analysis, syntactic/semantic feature analysis, SM/SFA/SSFA combination, and visual display) and subject area (English, Science, and Social Studies). Specifically, when comparing content areas Dexter and Hughes (2011) reported the largest effect sizes in Science ($ES = 1.05$) and the lowest effect sizes for math ($ES = .59$). With regard to differential effects by type of graphic organizer Dexter and Hughes (2011) reported a range of posttest effects (e.g., .74-1.2) for different kinds of graphic organizers, excluding visual displays. Similarly, Dexter et al. (2011) reported differential effects by type of graphic organizer on Science performance and concluded that all of the types of graphic organizers included in their analyses had large effects on Science performance.

It is important to note that while Dexter and Hughes (2011) and Dexter et al. (2011) reported differential effects sizes, they did not actually run any subgroup moderator analyses in their meta-analyses. In addition, their analysis of differential effects by type of graphic organizer lacked specificity, as it is impossible to compare individual effects of type of graphic from their
analysis. Thus, a major limitation of both the Dexter and Hughes (2011) and Dexter et al. (2011) analyses is the lack of moderator analysis.

Another limitation of the Dexter and Hughes (2011) and Dexter et al. (2011) meta-analyses was that they were unable to use contemporary analytic methods, specifically Robust Variance Estimation (RVE), in their analysis. This is important as they reported multiple effect sizes from the same studies, thus violating the assumption of independence. As a result, their analysis contained correlated effect sizes.

**Purposes for this Current Meta-Analysis**

The primary purpose of this meta-analysis was to extend the previous meta-analysis literature, as well as address several limitations of previous syntheses and meta-analyses of graphic organizer interventions. A secondary purpose was to describe implications for practice.

**Extending the Literature**

First, this meta-analysis updated previous meta-analyses and syntheses by synthesizing research through August 2019. This date range extended the LD graphic organizer research another 15 years (Kim et al., 2004); the LD computer-assisted graphic organizer research another six years (Ciullo & Reutebuch, 2013); the LD Science graphic organizer research another eight years (Dexter et al., 2011); and the graphic organizer for LD students research another eight years (Dexter & Hughes, 2011). Second, this meta-analysis extended the literature by including students with diverse characteristics (i.e., LD, ID, ASD, low achieving students, and students identified as general education [GE]) who represent the diverse populations that define contemporary inclusive classrooms.

**Methodological Approach**
This meta-analysis used contemporary statistical methods, such as Robust Variance Estimation (RVE) and between-case standardized mean difference (BCSMD). Calculating BCSMD enabled the inclusion of single-subject research. The use of RVE enabled the inclusion of studies containing multiple measures and/or more than one control group or treatment group. Without RVE, other methods of excluding effect sizes (ES) have to be used, as the assumption of independent effect sizes is violated when multiple ES are nested within individual studies. It is important to note that while Dexter and Hughes (2011) and Dexter et al. (2011) did not use RVE in their meta-analysis, a deliberate decision was made to extend their work and not both replicate and extend this work for this meta-analysis. The rationale is as follows. First, many of the studies included in Dexter and Hughes (2011) and Dexter et al. (2011) did not contain reading as part of the intervention, which is a criterion for this meta-analysis, or did not include measures of content knowledge (Bos & Anders, 1990; Darch et al., 1986; Darch & Eaves, 1986). Second, because the results of many of the included studies were analyzed using ANCOVA or MANCOVA, there was no pre-test equivalence data, which was also a criterion for inclusion for this meta-analysis (Hudson, 1996; Reyes et al., 1989). Third, some studies used advanced organizers which are outlines and do not meet the operational definition of a graphic organizer for this meta-analysis (Darch & Gersten, 1986). Fourth, even if the eleven Science and Social Studies graphic organizers included in the Dexter and Hughes (2011) and Dexter et al. (2011) studies had met all inclusion criteria, there would still not be enough studies \((k = 27)\) to run an RVE meta-regression analysis, as the recommended threshold is 40 studies (Tanner-Smith & Tipton, 2013).

**Moderator Analysis**
This meta-analysis also extended the graphic organizer literature by including a moderator analysis of variables that may impact the effectiveness of graphic organizers. While previous syntheses and meta-analyses on graphic organizers have described effect sizes or reported differential effects, no statistical analyses of the variables that moderate graphic organizer interventions have been computed.

**Study Design.** Since this is the first meta-analysis to include both group design and single-subject studies, study design was included as a moderator.

**Participant Type.** One goal of this meta-analysis was to represent the diversity of students in contemporary classrooms. Thus, students with a wide range of ability were included in this meta-analysis: general education, low achieving students (LA), students on the autism spectrum (ASD), students with intellectual disabilities (ID), and students with learning disabilities (LD).

**Grade.** To refine the descriptive analyses completed by Kim et al. (2004), Dexter and Hughes (2011), and Dexter et al. (2011), grade (i.e., upper elementary, middle, and high school) was included as a moderator in this study.

**Subject.** Science and Social Studies were chosen as moderators to provide an empirical analysis of the descriptive analysis completed by Kim et al. (2004).

**Graphic Organizer Type.** Furthermore, graphic organizer type was chosen as a moderator for two reasons. First, modal graphic organizers are most prevalent in instructional and teaching material, however they have not been independently represented as a graphic organizer type in prior syntheses and meta-analyses (Dexter & Hughes, 2011; Dexter et al., 2011; Kim et al., 2004). In addition, this meta-analysis introduced the disciplinary graphic
organizer as a new type to reflect contemporary views of knowledge representation within the disciplines.

**Type of Intervention.** In addition, this meta-analysis provided a more refined look at how instructional features such as the inclusion of reading instruction within a text-based graphic organizer intervention moderated treatment effects. This is an important moderator given that lower level cognitive processes of reading impact higher levels of processing. Thus, if students receive reading instruction in tandem with graphic organizer instruction, the impact of these lower level processes may be reduced.

**Teacher Instruction.** Teacher instruction was included as a moderator, as previously it has been reported to be a statistically significant moderator of intervention effectiveness for writing (Gillespie & Graham, 2014). Teacher instruction was defined as instruction that was teacher-led at the beginning of the intervention and where students either observed the teacher constructing a graphic organizer using instructional techniques such as explicit instruction, or the instructor led a modified think-aloud where the instructor created the graphic organizer and the students discussed the decisions being made as they created the graphic organizer in tandem with the teacher. Teacher instruction did not include any instruction on general graphic organizer creation prior to instruction.

**Consideration of Quality Indicators**

Finally, no previous graphic organizer meta-analysis has been conducted with consideration of the quality indicators outlined by Horner et al. (2005) and Gersten et al. (2005) and the What Works Clearinghouse Quality Standards for Group Design and Single Subject Studies (Kratochwill et al., 2010). Thus, this meta-analysis included a descriptive review and analysis of all studies using these indicators.
Research Questions

Specifically, the following research questions were examined: (1) What are the effects of text-based graphic organizer interventions on content learning for students in grades 3-12? and (2) What study, participant, graphic organizer, and intervention characteristics moderated treatment effects?
CHAPTER TWO

METHOD

Study Inclusion Criteria

To be eligible for this review, studies had to meet the following criteria: (a) studies were published in English (including all intervention materials) in a peer-reviewed journal or as part of a dissertation between January 2000 and August 2019; (b) participants were students in grades 3-12 identified as general education (GE), low achieving (LA), learning disabled (LD), Autism Spectrum Disorder (ASD), or intellectually disabled (ID); (c) studies involved a true experimental design with randomization, a quasi-experimental design with a treatment-control group comparison and reported pre-test equivalence before intervention, or a single-subject design; (d) studies included at least one Science or Social Studies graphic organizer intervention, and the graphic organizer was used to show conceptual relationships discussed in the text or was used to outline or organize this conceptual information in the text, compared to: another non-GO reading intervention or business-as-usual condition; (e) studies that only included interventions that involved graphic organizers that physically represented either the hierarchical, rhetorical/modal, or disciplinary structure and relationships of either the disciplines of Science or Social Studies or Science and Social Studies texts were included; and (f) studies contained at least one measure of Science or Social Studies content knowledge.

Search Strategies

This review used a three-part search: electronic, hand, and ancestral search. Figure 1 provides a PRISMA chart to summarize this search process. First, previous syntheses and meta-analyses on graphic organizers were reviewed to identify search terms (Ciullo & Reutebuch, 2013; Dexter & Hughes, 2011; Dexter et al., 2011; Kim et al., 2004; Nesbit & Adesope, 2006;
Schroder et al., 2017). As a result, the following search terms were identified: *concept map* OR *cognitive map* OR *cognitive organizer* OR *graphic organizer* OR *knowledge map* OR *node-link map* OR *visual display* OR *content enhancement* OR *text structure* OR *semantic map* OR *semantic feature analysis* OR *Venn diagram* OR *flow chart* OR *story map* OR *semantic organizer*. An electronic search was first conducted using these key terms and the following databases: Academic Search Premier, ERIC, Education Source, PsychINFO, SOCINDEX with Full Text, using the following limiters: peer review, published in English, published between 2000 and August 2019. To extend the initial search, a second electronic search was conducted in August 2019, using ProQuest Dissertation/Theses and the same search terms listed above. In addition, this search included the following limiters: English and after 1999. Using this procedure, 21,329 articles, dissertations, and theses were identified.

Next, articles were screened, and eligibility was determined using the PRISMA eligibility and selection procedure (Moher et al., 2009), as outlined in Figure 1. To start this process, articles were entered into Mendeley, a reference software management tool. Duplicate articles ($k = 5,648$) were deleted both automatically by the Mendeley database and by the author, and 15,681 unique abstracts remained.

Once duplicates were removed, articles were screened by the author and 15,189 abstracts were excluded because they did not meet basic inclusion criteria. Of these 15,189 screened abstracts, 14,191 were excluded as they did not cover educational topics and included topics such as proteins and cartography. Fifteen articles were excluded because they were not published in English. Twenty-eight articles were meta-analyses or synthesizes. Two hundred and seventy-one articles did not contain a graphic organizer intervention, and 135 studies were excluded because they did not contain a Science or Social Studies measure of content knowledge. Three hundred
and thirty-four screened articles were excluded as they were non-experimental, and 215 articles were excluded because the study included post-secondary students, graduate students, or teachers.

Once abstracts were screened, 492 full text articles were analyzed to determine eligibility. Seventy-six studies were determined as ineligible for this review because of the participants. Twenty-five were excluded because they were non-experimental. Seventy-seven studies were excluded because of study design. Examples of excluded studies were those that employed alternating treatment single-case design (Reed et al., 2016; Williams et al., 2005) or group design studies that compared one graphic organizer intervention to another (Brown, 2003; Odom & Kelly, 2001; Royer & Royer, 2004)

One hundred and sixty-one articles were excluded because they did not contain a Science or Social Studies graphic organizer in the intervention, or because the graphic organizer included in the intervention did not provide a physical representation of the relationships between ideas, or because there was no read material in the intervention. Studies that were deemed ineligible due to the characteristics of the graphic organizer included: those studies that used Wh-Question graphic organizers, question exploration routines, or self-questioning strategies, or studies that used only a vocabulary based graphic organizer (Bulgren et al., 2014; Bulgren et al., 2011; Wood, Browder, & Flynn, 2015; Bethune & Wood, 2013). Studies that only included oral presentations of Science and Social Studies content through lecture and discussion were also excluded (Browder et al., 2013; Odom & Kelly, 2001). One hundred and thirty-two studies that did not include a measure of Science or Social Studies content knowledge or only contained a general measure of reading comprehension (i.e., normed comprehension assessments, or assessments that asked for literal or inferential recall of details from the reading, as well as those
that included questions about main ideas and supporting details) were excluded (Wijekumar, et al., 2018). Two studies were excluded because they were meta-analyses, and two studies were excluded because the instructional materials used in the intervention were not published in English. Finally, one study was excluded because there was insufficient data to calculate an effect size.

In addition to an electronic search, the following journals were hand searched from 2000-August 2019: *Elementary School Journal, Exceptional Children, Focus on Autism and Other Developmental Disorders, Journal of Behavior Disorders, Journal of Educational Psychology, Journal of Learning Disabilities, Journal of Special Education, Middle School Journal, National Council of Social Studies Journal, Reading and Writing, Reading Research Quarterly, Remedial and Special Education, Research and Science Education, Review of Educational Research, Science Education*. These journals were chosen as they represent leading journals in Science and Social Studies education, reading research, and research on students with disabilities. From this hand search, no unique articles that met the inclusion criteria were identified.

Finally, the reference pages of meta-analyses and syntheses on graphic organizer and Science/Social Studies interventions conducted between 2000 and 2019 (Ciullo & Reutebuch, 2013; Dexter et al., 2011; Dexter & Hughes 2011; Kim et al., 2004; Nesbit & Adesope, 2006; Schroder et al., 2017), as well as all identified articles, were reviewed to identify additional articles. In addition, the reference lists of recent meta-analyses and reviews on content learning and reading comprehension were screened (Ciullo et al., 2016; Hebert et al., 2016; Scruggs et al., 2010; Knight & Sartini, 2015; Swanson et al., 2014; Wanzek et al., 2010). No unique articles that met the inclusion criteria were located during this process.
As a result of this search process, which ended in August 2019, 16 articles and
dissertations that met inclusion criteria were identified (Arslan, 2006; Berry et al. 2013; Boon,
Burke, Fore, & Hagan-Burke, 2006a; Boon, Burke, Fore, & Spencer, 2006b; Carnahan &
Williamson, 2013; Carnahan et al., 2016; Cash 2013; Ciullo, Falcomata, Pfannenstiel, &
Billingsley, 2015a; Ciullo, Falcomata, & Vaughn, 2015b; DiCecco & Gleason, 2002; Guastello
et al., 2000; Ledger, 2003; Nair & Narayanasamy, 2017; Schenning et al., 2013; Tastan et al.,
2008; Zakas et al., 2013).

Study Coding

The coding manual for this study was modified from a coding manual from a previous
writing intervention meta-analysis (Gillespie & Graham, 2014). Studies were coded for the
following variables: study level variables (i.e. author, year of publication, publication type,
experimental design, single-subject design, unit of assignment, pre-test equivalence, locale);
participant characteristics (i.e., grade, participant disability category, identification of disability,
pre-study fluency, pre-study comprehension); intervention qualities (i.e. content area, subject,
treatment and comparison instruction location, person delivering intervention, teachers per
condition, teacher training, baseline instruction, content area instruction in intervention);
intervention and control group descriptions; intervention intensity (i.e., size of intervention
group, duration, frequency, dosage); text (i.e. text description, text level, text type, text
modifications, text reader); graphic organizer (i.e., graphic organizer type, graphic organizer
picture, mapper, type of response); measure (i.e. measure names, question type, measure type,
form of student response, measure support, measure reliability, frequency of measure, time after
intervention); information needed to calculate effect sizes (i.e. number of students in treatment
and control conditions, pre-test and post-test mean score of content learning for each group); and
study quality indicators (i.e. teacher effects, fidelity of implementation, group design- random assignment, attrition data; single subject-independent variable manipulation, inter-assessor frequency, number of raters and statistics, data display, number of phases, and number of data points).

**Inter-Rater Reliability**

The author coded all articles. In addition, thirty percent of included articles \( (k = 5) \), were randomly chosen using an online digitizer, and then coded by a second coder (a co-chair of this dissertation). Percentage of exact agreement was 89%. This meets the threshold for sufficient agreement (NAEP, 2013). All discrepancies in coding were resolved through discussion and consensus.

**Statistical Analysis**

*Standardized Mean Difference Effect Sizes (ES)*

\( ESs \) were calculated for all measures of content knowledge in individual studies. This meta-analysis included 18 \( ESs \) in 16 studies. The standard mean difference for group design studies, Cohen’s \( d \), was calculated as the post-test mean score of content learning for the control group (the non-GO intervention group) subtracted from the posttest mean score of content learning from the treatment (GO-intervention group), and then divided by the pooled standard deviation of both groups. The standardized mean differences for all studies were computed using the statistical software STATA 16.0 and double checked using Wilson’s Practical Meta-Analysis Effect Size Calculator (Wilson, n.d.). Finally, all standardized mean difference effect sizes were corrected for small sample bias, using the Hedges’ \( g \) correction in STATA 16.0 (Hedges & Olkin, 1985).

*Between-Case Standardized Mean Difference Effect Sizes*
For single-subject studies, the between-case standardized mean difference (BCSMD) was calculated using the procedures outlined by Zelinsky and Shadish (2018) and Valentine et al. (2016). First, the program UnGraph was used to digitize individual data points from the graphs of the single-subject studies. Then this data was reviewed and prepared, and the between-case standardized mean difference was computed using the SCHDLIM application (Pustejovsky, 2016). Finally, as recommended by Zelinsky and Shadish (2018), all between-case standardized mean difference effect sizes were corrected for small sample bias, using the Hedges’ $g$ correction in STATA (Hedges & Olkin, 1985). Hedges’ $g$ was also manually computed for all between-case standardized mean differences to ensure computational accuracy.

**Sensitivity Analysis**

Using STATA 16.0, a forest plot was created and standardized mean differences were examined for outliers using Hoaglin et al.’s (1986) definition of an outlier, any effect size or sample size that falls three interquartile ranges above or below the mean (i.e., above 75th percentile or below the 25th percentile). Through this process, it was observed that the effect size for Guastello et al. (2000) fell above the 75th percentile ($g = 5.96$) and thus it was winsorized to the 75th interquartile ($g = 4.87$).

**Analysis of Heterogeneity**

Heterogeneity analysis was conducted using STATA 16.0. The statistics from the tests for heterogeneity were analyzed (including $t^2$ and $I^2$) to determine if effect sizes differ by significantly more than what is expected by chance alone.

**Calculation of Overall Mean Effect Across Studies**

A random-effects model with robust variance estimation (RVE) was used to calculate the overall mean effect across studies, as variation across studies was expected. Specifically, a
correlated effects RVE approach was applied given that four effect sizes were included from two primary studies, with effect sizes being nested within studies and violating the assumption of independent effect sizes. RVE is a recommended approach with correlated effect sizes, as it addresses heterogeneity by adjusting standard errors (Tanner-Smith & Tipton, 2013). In addition, RVE was used, as Pustejovsky et al. (2014) recommend using RVE when incorporating BCSMD effect sizes into meta-analysis, as the small number of participants may lead to understated standard errors (Valentine et al., 2016).

The procedures outlined by Tanner-Smith and Tipton (2013) along with STATA 16.0 were used to compute the average weighted effect across studies. Per recommendations by Tanner-Smith and Tipton (2013) a within-study correlation between effect sizes of ($\rho = .8$) was used. This value was determined to be appropriate after a sensitivity analysis for rho values between .1 to .9 was run per the guidelines outlined by Tanner-Smith and Tipton (2013). In addition, 95% prediction intervals of true effect sizes were calculated using ($\mu - 1.96\sqrt{\tau^2}$) in tandem with the guidelines presented by Borenstein et al. (2009). Because these statistics indicated variation between studies, a moderator analysis was conducted to account for the significant variation using meta-regression.

**Moderator Analysis**

Because this meta-analysis contained 16 studies, meta-regression was an appropriate analytic method, as Borenstein et al. (2009) state the threshold for using meta-regression is the inclusion of ten or more studies to predict heterogeneity in effect sizes. The following moderators were determined a priori to be analyzed: study type (i.e. group design, single-subject); grade (i.e., elementary, middle, high school); subject (i.e., Science, Social Studies, Science & Social Studies); intervention type (i.e., graphic organizer only, graphic organizer and
reading instruction); intervention practices (i.e. teacher instruction versus no teacher instruction); graphic organizer type (i.e., concept map, disciplinary map, or modal map); student type (i.e., no identification of learning or reading risk, mixed population, general education and low achieving or learning disabled students, students at risk [i.e., LD, ID, ASD]). While robust variance estimation (RVE) was used to calculate the average weighted effect sizes across studies, RVE was not used for meta-regression, as a minimum of forty studies is recommended for sufficient statistical power (Tanner-Smith & Tipton, 2013).

STATA 16.0 was used to recode all moderators as dummy variables. Prior to running the planned meta-regression, a correlation matrix was created and examined to ensure there were no correlated variables. Upon examination it was observed that there was collinearity between student type and study type; student type and graphic organizer type; instruction and subject; and intervention type and graphic organizer type. Thus, subject, graphic organizer type, and student type were dropped from the model. Specifically, student type and graphic organizer type were dropped from the model because they were correlated with more than one variable, and subject was dropped from the model because it was assumed that this would not be a significant moderator.

**Publication Bias**

The potential for publication bias was examined using multiple methods. First, funnel plots were generated in STATA 16.0 and were then inspected to determine the distributions of studies with regards to effect sizes and standard errors. Next, an Egger’s test for small study effects was run in STATA 16.0 to statistically evaluate the presence of publication bias. Last, a non-parametric Trim and Fill analysis was run in STATA 16.0 to estimate the true effect size and an estimate of missing studies.
Descriptive Analysis of Quality Indicators

To evaluate the quality of studies using the What Works Clearinghouse criteria, articles were divided by group design or single-subject design, and individual quality indicators were coded using the version 3.0 Standards Handbook for group design and the Version 1.0 Single-Case Design Technical Documentation document (Kratochwill et al., 2010; WWC, 2014). The criteria in the Determinants of a WWC Study Rating decision chart in the group design technical document were applied to evaluated group design studies, and studies were evaluated for: random assignment, pre-test equivalence, amount of attrition, and percentage of attrition.

For the single subject studies, the standards outlined in the Criteria for Designs that Meet Evidence Standards were applied (Kratochwill et al., 2010). Studies were specifically evaluated with regard to: systematic manipulation of the intervention; interassessor measurement and agreement (i.e., 20% of data points in each phase need to be measured by more than one assessor with minimal interobserver agreement thresholds being met); three attempts to demonstrate an intervention effect; a minimum of three or five data points per phase (number of phases varies for different types of single-subject designs) with a set number of phases.

It is important to note that the standards outlined in the version 4.0 handbook were not applied, as that version was published in 2017, which is after all of the studies in this review but one were published. Once individual criteria were coded for each study, the WWC criteria were applied to categorize studies as: meets standards, meets standards with reservations, does not meet standards. Study quality was treated descriptively and was not included as a moderator in this meta-analysis, following the recommendation by Valentine (2019).

In addition to the quality indicators outlined by the WWC, other indicators of study quality were evaluated. Specifically, measure reliability and treatment fidelity were evaluated.
CHAPTER THREE

RESULTS

Before answering research questions in this meta-analysis, a brief overview of included studies is provided. Table 1 provides an overview of the study characteristics, assessments, and study quality. Table 2 provides an overview of the materials and intervention components.

Description of Studies

Study Characteristics

Sixteen studies which were published between the years of 2000-2015 were included in this meta-analysis. Two studies were dissertations \((k = 2)\) and the remaining studies \((k = 14)\) were published in peer-reviewed journals. Study design ranged from experimental \((k = 4)\), to quasi-experimental \((k = 6)\), to single-subject \((k = 6)\). Studies were conducted in three countries including, Malaysia \((k = 1)\), Turkey \((k = 2)\), and the United States \((k = 13)\).

Subject

Studies included Science \((k = 7)\), Social Studies \((k = 8)\), and mixed Science and Social Studies interventions \((k = 1)\). For the seven Science interventions, four studies included Biology as a sub-discipline, one study included Astronomy, one study included Geology, and in one study, the sub-discipline was undefined. For the eight Studies that included Social Studies interventions, seven studies included History and for one study, the sub-discipline was undefined. The study that included Science and Social Studies did not identify the sub-discipline.

Participants

Studies included students in upper elementary \((k = 4)\), middle \((k = 8)\), and high school \((k = 4)\). Studies included students with and without disabilities. Four studies included students
identified as general education with no identified disabilities. Four studies included students that were in mixed population classes (i.e., GE/LD or GE/LA). One study included low-achieving students. Two studies included students identified with learning disabilities. One study included students identified as LD or ID. Two studies included students identified as ASD and two studies included students identified as both ASD and ID or only ASD.

**Intervention Delivery and Teacher Training**

Five studies were delivered in a general education classroom, three studies were delivered in a resource room, one study was conducted in a special education classroom, and there were seven studies where the location of the intervention could not be determined. Interventions were delivered by general education teachers \( (k = 5) \), special education teachers \( (k = 3) \), a combination of general education and special education teachers \( (k = 2) \), the researcher or member of the research team \( (k = 3) \), and in three studies, the person implementing the intervention was not specifically identified.

Implementor training varied by type, level of training, and length of training. In five studies teachers received direct intervention training. For instructors who received direct intervention training, training sessions varied from 45 minutes \( (k = 1) \) to two and a half to three hours \( (k = 2) \) to two training sessions \( (k = 1) \) to six weeks \( (k = 1) \). In three studies, instructors did not receive direct intervention training. Rather, in two of these studies, instructors received 45 minutes of training on Inspiration software and in another study, teacher coursework in a degree program was counted as training. Eight studies did not report intervention training.

**Instruction**

Five studies included interventions that were only graphic organizers, while 11 studies included both graphic organizers and reading instruction (i.e., graphic organizer intervention
with fluency, vocabulary, and/or text structure instruction or activities). Teacher instruction (e.g., teacher modeling, think-alouds, or explicit instruction on the cognitive processes of completing a graphic organizer) prior to students completing their graphic organizers in a group was reported in eight studies.

**Instructional Intensity and Dosage**

In 10 studies, students received intervention in a whole group format with group size reported between 20 and 48 students. Students in six studies received intervention in a small group, 1:1, or both small group and 1:1 formats.

Instructional duration varied from eight days ($k = 2$), to three to six weeks ($k = 7$), to an entire semester, two quarters, or 12 weeks ($k = 3$). Intervention duration was not reported in four studies. Frequency of session data was limited, with only six studies reporting this information. For those studies reporting this information, frequency of intervention varied from three to five sessions per week. Intervention dosage data or session length was also limited, as only four studies reported this information, with session lengths varying from 40-90 minutes.

**Description of Materials**

**Text.** Researcher-generated and academic texts were used in the graphic organizer interventions. Academic texts used in the studies included: textbooks ($k = 8$), course readers ($k = 1$), trade books ($k = 1$), basal readers ($k = 1$), and textbooks along with trade books ($k = 1$). Researcher-generated texts were included in four studies. These included texts with reduced language and picture cues for students with ASD and conceptual change texts in one study that included GE students without identified disabilities. Students were matched to text using readability formulas in three studies. One study controlled for passage length and reading level but did not report how this was achieved. Six studies matched readers to text by grade level and
the readability of text was not reported in six studies. In addition, it was reported that teachers read texts to students in four studies, students read texts in seven studies, and it could not be determined who read the text in five studies.

**Graphic Organizer.** Three types of graphic organizers were included in these studies. Five studies included concept maps, 10 studies included modal graphic organizers, and one study included disciplinary graphic organizers. Partially completed graphic organizers were given to students in five studies, templates were given to students in four studies, templates and partially completed graphic organizers were given to students in two studies, students completed graphic organizers from scratch in three studies, and in two studies, the initial form students started with when creating graphic organizers was not reported. Students completed the graphic organizer in 15 studies and the teacher completed the graphic organizer in one study. In one study, computer-generated graphic organizers were used, in two studies both computer and pencil and pen graphic organizers were used, in one study students provided oral responses to their teacher, and in one study student placed their responses on the graphic organizer; however, the majority of studies ($k = 11$) included pencil and paper graphic organizers.

**Description of Measures**

Researcher-generated measures were most often used in studies ($k = 12$), with publisher-modified or publisher-created measures being used in one study. Three studies did not report who generated the measure. Question formats included: structured question ($k = 1$); number of relational knowledge statements ($k = 1$); correct placement of information on a graphic organizer ($k = 1$); multiple choice ($k = 4$); open ended production tasks or short answer ($k = 5$); and multiple choice and short answer, or multiple choice and correct placement ($k = 4$). For studies that reported how students responded to assessment items, student responses varied between
correct placement of answers \((k = 1)\), to providing oral answers \((k = 3)\), to writing answers \((k = 4)\), to providing both oral and written answers \((k = 1)\). Students also received varying levels of support when answering questions including: no reported support \((k = 10)\), rereading of text \((k = 1)\), look backs and error correction \((k = 2)\), and assessor transcription \((k = 3)\).

**Descriptive Analysis of Quality Indicators**

There was a variety of study quality when the 2013 WWC Quality Standards (Kratochwill et al., 2010; WWC, 2014) were applied to the studies in this meta-analysis. Table 1 outlines the overall WWC Quality Standard designation for each study.

Of the 10 group design studies in this meta-analysis, only one study met the criteria for random assignment, pre-test equivalence, and equal attrition. Six studies met the standards with reservation because they did not include random assignment. Three studies met the standards with reservation because they did not report attrition information, thus it was impossible to determine if attrition was equal.

For the six single-subject studies, four studies met the first three standards for systematic manipulation of the independent variable, systematic measurement over time by more than one assessor, and appropriate attempts to show an intervention effect. However, these studies were categorized as “meets standards with reservation” because the last standard of having five data points per phase was not met. Each of these studies had a minimum of three data points. Two single subject studies did not meet the second standard of IOA collection and reporting, thus resulting in a “does not meet the standard” quality indicator classification. It is important to note that both of these studies reported that data was collected in over 20% of data items, the inter-assessor agreement met minimum thresholds, and the dependent variable was measured over time by multiple assessors. However, these studies failed to specify that the percentage of data
collected was in each phase, which is why they were categorized as “did not meet.” Thus, it is highly likely that this is an issue with reporting specificity and that these studies may have actually met the single subject standards with reservations.

Fidelity of instruction was only reported in half of the studies included in this review ($k = 8$). Of those studies, interrater reliability was only reported in four studies and varied from 99.6-100%. Formal technical adequacy scores for assessments were not reported for any of the single subject studies included in this meta-analysis ($k = 6$). For the group design studies, technical adequacy scores were reported for five studies, with five studies not reporting technical adequacy. Of the five studies that did not report technical adequacy, two studies did report interrater reliability between two scorers (Boon et al., 2015a, Boon et al. 2015b).

**Research Question One: What are the effects of text-based graphic organizer interventions on content learning for students in grades 3-12?**

For RQ 1, 18 ESs were calculated from 16 studies. Table 1 and Figure 2 show an average effect size for each study. The intercept only RVE model indicated an overall, average weighted effect size of 1.65 [0.96, 2.35], indicating a statistically significant effect of graphic organizers on students’ content learning in Science and Social Studies. Specifically, on average, students who participated in a graphic organizer reading intervention in Science or Social Studies scored slightly more than one and one-half standard deviations higher than participants in control groups on measures of content learning.

Individual effect sizes ranged from 0.115 to 4.87. A $\tau^2$ of 1.245 provided evidence of variation between true effect sizes, resulting in a 95% prediction interval of -0.537 to 3.84. This indicates that 95% of true study average effect sizes found in the population should be between...
this range. An $I^2$ value for the model indicated that 94.24% of the observed variance was beyond sampling error and the result of true between-study variance.

**Research Question Two: What study, participant, graphic organizer, and intervention characteristics moderate treatment?**

Results of the multivariate meta-regression with the moderators of study type, grade level, intervention type, and inclusion of instruction can be found in Table 2. There was evidence of collinearity between the following variables: student type and study type ($r = .71$); student type and graphic organizer type ($r = .80$); instruction and subject ($r = .742$); intervention type and graphic organizer type ($r = .73$), so student type, graphic organizer type, and subject were dropped from the model. A multivariate meta-regression model that included the following variables: study type, grade level, intervention type, and inclusion of instruction explained 23% of the true between-study variance. The $Q_{res}$ statistic ($Q = 145.65, p < .000$) indicated a statistically significant amount of unexplained heterogeneity in the model. Teacher instruction was a statistically significant moderator ($\beta = 1.445, p < 0.05$). An $I^2$ value of 92.21% indicated a large amount of the residual variation was due to true heterogeneity between studies.

**Publication Bias**

Inspection of the funnel plot showed relative symmetry (see Figure 3). However, visual analysis did show the absence of studies with smaller $ES$’s and larger standard errors. Results of the Egger’s test, $B_I = 2.59$, ($SE = 1.325$), $z = 1.95$, $p = .051$, indicated the possible presence of publication bias. However, results from the non-parametric trim and fill test were identical for observed ($k = 18$) and observed and imputed ($k = 0$) studies, $ES = 1.685$ [1.074, 2.296].
CHAPTER FOUR

DISCUSSION

In this meta-analysis, 16 studies with 18 effect sizes were reviewed to evaluate the effects of text-based graphic organizers on learning Science and Social Studies. This type of meta-analysis is important given that students are demonstrating low levels of content knowledge on national measures starting in upper elementary grades, and reading is the primary way students acquire this knowledge (NAEP, 2015; NAEP, 2014; Swanson et al., 2014).

Overall, this meta-analysis was motivated by a unique conceptual framework that connects the cognitive processes involved in reading, learning, and the creation of graphic organizers to explain why text-based graphic organizer interventions should have a positive effect on content learning. This theoretical view was shaped by the cognitive views of reading and learning which suggest that reading to learn is effective when a coherent mental representation of read information is created, as the result of higher and lower levels of cognitive processing of text (van den Broek et al., 2017). The conceptual framework for this meta-analysis also borrows from the cognitive processing view of graphic organizers which suggests graphic organizers help individuals effectively engage in the highest levels of cognitive processing (i.e., selection, integration, and organization) to create coherent mental representations of text (McCrudden & Rapp, 2017).

This conceptual framework also informed the selection of key moderators. The selection of multiple moderators was largely informed by the Cognitive View of Reading, which theorizes that lower levels of text processing (i.e., vocabulary and fluency) influence higher levels of text processing (van den Broek et al., 2017). This theory also suggests that cognitive processing of
text is influenced by individual characteristics, such as disability status and age (van den Broek et al., 2017). Thus, participant type and grade were selected as moderators.

The inclusion of another moderator, graphic organizer type, was informed by contemporary views of knowledge representation, including Schema Theory and Expertise Theory (Chi, 2006; Chi & Ohlsson, 2005). Since graphic organizers represent mental representations of learned and read knowledge, graphic organizer type was included as a moderator to examine if there were differences in content learning between students who used modal graphic organizers that organized information around the modes of a text, students who used concept map graphic organizers that represented information using semantic links, and students who used disciplinary graphic organizers that represented information by key theories. 

**Research Question One: What are the effects of text-based graphic organizer interventions on content learning for students in grades 3-12?**

Across the studies included in this meta-analysis, graphic organizers facilitated students’ learning Social Studies and Science content through reading with an overall $ES$ of 1.65. This is considered a large effect according to Cohen (1988). Graphic organizers positively impacted content learning in both Science and Social Studies for a range of students, including students with varying levels of reading and disability designations, across range of grades, across a range of study types, and with a range of texts.

This evidence supports the theory that graphic organizers facilitate the highest levels of cognitive processing while reading to learn, and graphic organizers enable even students with disabilities or low achievement, who particularly struggle with selecting, attending, and organizing text, to learn while reading. This meta-analysis also shows that with appropriate
considerations of text and instruction, graphic organizers aid students in learning from relatively complicated academic texts, including Science and Social Studies textbooks.

This $ES$ converges with previous reviews (Dexter et al., 2011; Dexter & Hughes, 2011) that reported large overall weighted effect of graphic organizers on content learning ($ES = 1.052$ and $ES = .91$, respectively). This meta-analysis contributes to the field, as it extends the most frequently cited meta-analyses of graphic organizer interventions in the content areas by Dexter and Hughes (2011) and Dexter et al. (2011) and continues to demonstrate that graphic organizers not only support students with LD in learning from content area texts, but it extends the literature by demonstrating graphic organizers can also support other struggling students, even those who do not have an identified disability, such as students who are low-achieving in reading, or students identified with disabilities, including students identified as ASD or ID.

This meta-analysis also extends the literature on graphic organizer research by using contemporary analytical and statistical methods. First, methods were used to allow single subject studies and group design studies to be included in a graphic organizer meta-analysis for the first time. Second, this meta-analysis used a random effects model, in contrast to a previous meta-analysis on content learning for students identified as LD, which used a fixed effect model (Dexter & Hughes, 2011). In contrast to a fixed effect model, which assumes that differences in effect sizes is only the result of sampling error, a random effects model assumes that the true effect size will be different across studies as a result of differences within each individual study (Borenstein et al., 2009). Thus, given the studies in this meta-analysis included different participants, as well as differences in how interventions were implemented, a random effects model is a better model for this data. Third, $RVE$ was used to treat dependent effect sizes
appropriately and reduce bias when weighting studies. Use of these advanced and up-to-date methods ensures confidence in the findings reported in this paper.

**Research Question Two: What study, participant, graphic organizer, and intervention characteristics moderate treatment?**

Given that student type was correlated with study type as well as graphic organizer type, subject was correlated with instruction, and graphic organizer type was correlated with intervention type, these three moderators were dropped from the analysis. As a result of collinearity, these moderators could not be analyzed. The moderators of study type, grade level, and intervention type, which were selected a priori, were not statistically significant and did not predict a significant amount of variance in effect sizes. Notable differences between studies with regard to participants, intervention, and the texts used may explain why these were not significant moderators.

Instruction made a positive and significant contribution to predicting $ES$ \(\beta = 1.445, p = .027\). This finding indicates that text-based graphic organizer interventions were more effective when students received teacher instruction on how to complete the graphic organizer at the beginning of an intervention. While Dexter and Hughes (2011) suggested that students need explicit instruction in graphic organizer interventions and that effective instructional practices, such as modeling and providing feedback, positively impacted the effectiveness of graphic organizer interventions on student learning, this was not empirically or directly examined in their study. Thus, the present meta-analysis shows that the use of effective instructional practices does predict a significant amount of variance and moderated treatment effects.
Limitations

Considerations should be made when interpreting the results of this meta-analysis. First, study quality should also be considered when interpreting results. It is important to remember that of the 16 studies included in this meta-analysis, only one study fully met the Version 3 WWC standards (WWC, 2014) without reservation.

Another limitation of this meta-analysis was the outcome measures. Three studies did not describe the origin of the measures used and in the remaining thirteen studies, the measures were researcher-generated or publisher created and technical adequacy was not consistently provided \((k = 5)\). In addition, only one study used multiple measures to assess acquisition of content knowledge.

Another limitation of this meta-analysis is that while an exhaustive review of the literature was conducted, only peer-reviewed journal articles and dissertations accessible through search engines were included. Thus, it is possible that unpublished reports, dissertations not accessible through online search engines, data presented at conferences, or data in book chapters with potentially smaller or null effects were not located. Furthermore, while results from the non-parametric trim and fill test suggested that there were no missing studies with the best estimate of the unbiased effect size, visual analysis of funnel plots and results of Egger’s test provided some evidence of small study effects (i.e., smaller studies are missing), particularly for studies with smaller ES’s and larger standard errors. Collectively, these analyses suggest while publication bias may exist, it is most likely minimal.

Interpretation of the results from the moderator analysis of this meta-analysis should be made with the following considerations. First, the results of the meta-regression may have been limited by low statistical power given the relatively small number of effect sizes \((n = 18)\). It is
possible that there may have been greater statistical significance with the moderators if a larger number of studies had been included in this meta-analysis.

In addition, issues with collinearity greatly impacted the analysis of moderators. Specifically, participant type could not be analyzed as a moderator as it was a confounding variable in the meta-regression. This issue with collinearity was most likely caused by the breadth of participants included in this analysis and by the fact that general education or low achieving students were only included in group design studies with group instruction and students with autism or intellectual disabilities were only included in single-subject studies with small group and small group or individual instruction. Only students identified as learning disabled were included in both study designs.

Also, another limitation was that there was considerable heterogeneity in the model that was not explained through the chosen moderators for analysis. While this is not ideal, it is also not surprising given the variability in participants, intervention characteristics, texts, and graphic organizers.

The final limitation of the moderator analysis was that it was not possible to examine all three types of graphic organizers (i.e., modal, concept map, and disciplinary) because of issues with collinearity in the meta-regression. This moderator analysis of type of graphic organizer was further limited by the fact that only one study contained a disciplinary graphic organizer. Therefore, while this meta-analysis set out to redefine graphic organizers to align with contemporary theories of how knowledge is represented and analyze the types of graphic organizers being used in classrooms, this goal fell short as there was only one study and methodological limitations that prevented this analysis from happening.
Future Research

Future research would benefit from the following considerations, given the prevalence of confounding variables, largely as a result of a correlation between study type and participant type (e.g., participants without identified disabilities were only included in group design studies and participants identified as ASD were only included in single-subject studies). First, to understand the relationship between intervention size and graphic organizer effectiveness, group design studies including students with autism could be conducted. Second, a further meta-analysis looking specifically at students with LD could be conducted to better understand the instructional features that moderate the effectiveness of text-based graphic organizers on content learning. For example, to better inform instruction, the impact of group size or intervention intensity could be examined to compare the effect sizes of graphic organizer interventions on learning for students with LD who received intervention as part of a homogeneous whole group, as part of an inclusive classroom, or in a small group or 1:1 intervention.

In addition, given that methodological advances have allowed single-subject and group design studies to be compared in meta-analysis, consideration of factors that may prohibit students in single-subject studies to be treated as a group should be made if the intent is to include these types of studies in meta-analyses. Specifically, it is important to consider controlling for variability across participants, as including a breadth of students in a study who have a large range of disability designations, varying levels of participation in general education, varying reading levels, and varying levels of IQ, makes it impossible to treat these students as a comparable group. Thus, future single-subject research would benefit from the following suggestions. First, include only students with one disability type in a study (e.g., LD only and not
LD and ID). Second, control for other confounding factors between participants, such as IQ, reading level, inclusion into general education.

The breadth of this meta-analysis, including looking at multiple subjects and multiple types of participants, limited the depth of analysis that could be conducted. As more studies are published in this area, future graphic organizer meta-analyses could be narrower in scope and possibly focus on one subject (i.e., Science or Social Studies) or one or two population groups (i.e. general education and LD students, or just LD students, or only include students who are reading at grade level and participating in general education). This would allow for deeper analysis of the instructional practices that moderate graphic organizer intervention effectiveness.

This meta-analysis also set out to compare current designations of graphic organizers that better aligned with contemporary definitions of how knowledge is represented. This meta-analysis fell short of this goal, as the majority of text-based graphic organizers were organized around the dominant mode of the text and only one graphic organizer was organized around the dominant organization of the discipline. Given that graphic organizers act as cognitive facilitators and help students create mental representations of text, future research with learning and graphic organizers needs to be better grounded in contemporary theories of learning and cognition (specifically Expertise Theory), and it needs to examine if teaching students to organize and represent information around the dominant themes of a discipline is more effective than representing information using the modes of a text.

Finally, the moderator analysis in this meta-analysis may have been underpowered. Thus, to increase statistical power, additional studies on the effects of text-based graphic organizer interventions on learning should be conducted.
Implications for Practice

A secondary purpose for this meta-analysis was to inform practice. Given the results on national measures of content learning, the lower performance of students on these measures, and the continued expectation that students learn from complex text, it is essential that teachers have interventions that support the diverse needs of students in their classrooms to gain Science and Social Studies concept knowledge from reading. The findings from this meta-analyses show that generally speaking, graphic organizers facilitate learning from text with general education students, low achieving students, and students with LD, ID and ASD. In addition, because students who struggle to learn or read also often struggle with transferring strategies to novel situations, it is important to have strategies that can be deployed across multiple content areas. The findings from this meta-analyses support prior research that graphic organizers facilitate student learning from text in both Science and Social Studies.

However, while a large overall effect was found for the effectiveness of text-based graphic organizers on student learning in Science and Social Studies, it is important to remember that there was great variability in the interventions students received. Specifically, it is important to remember that general education or low achieving students only received whole group instruction, mostly participated in graphic organizer only interventions, and mostly read authentic texts. Students of mixed abilities (e.g., general education students and students with LD), all read authentic texts, received multi-component interventions that included reading, and largely received instruction on how to complete graphic organizers before completing graphic organizers on their own. Students with Autism received intervention in small groups or received individual instruction, mostly participated in interventions that included reading instruction, and read texts that had been heavily modified by the researcher to match their individual reading and
cognitive levels. Finally, LD students or struggling students all read academic texts, received multi-component interventions including reading, and almost all were provided with instruction at the start of their intervention. Thus, when thinking about implementing text-based graphic organizer interventions to aid students in reading to learn and acquire content knowledge, it is essential to consider students’ reading levels (including low-level processes such as vocabulary knowledge and fluency), instructional needs (including the need for modeling the process of completing graphic organizers prior to any group or individual work and supplementary reading instruction), and to make sure students are appropriately matched to text.
### Table 1

*Study Characteristics, Study Quality, and Average Effect Sizes*

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Design</th>
<th>Study Location</th>
<th>Publication</th>
<th>N</th>
<th>Subject</th>
<th>Assessment Type</th>
<th>Study Quality</th>
<th>Hedge’s g ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arslan (2006)</td>
<td>E</td>
<td>Turkey</td>
<td>J</td>
<td>N = 135</td>
<td>S</td>
<td>CD</td>
<td>M</td>
<td>0.115</td>
</tr>
<tr>
<td>Berry et al. (2013)</td>
<td>Q</td>
<td>US</td>
<td>J</td>
<td>N = 48</td>
<td>S</td>
<td>RG</td>
<td>MWR</td>
<td>0.919</td>
</tr>
<tr>
<td>Boon et al. (2006a)</td>
<td>Q</td>
<td>US</td>
<td>J</td>
<td>N = 44</td>
<td>SS</td>
<td>CD</td>
<td>MWR</td>
<td>1.99</td>
</tr>
<tr>
<td>Boon et al. (2006b)</td>
<td>E</td>
<td>US</td>
<td>J</td>
<td>N = 49</td>
<td>SS</td>
<td>CD</td>
<td>MWR</td>
<td>1.793</td>
</tr>
<tr>
<td>Carnahan &amp; Williamson (2013)</td>
<td>SC</td>
<td>US</td>
<td>J</td>
<td>N = 3</td>
<td>S</td>
<td>RG</td>
<td>MWR</td>
<td>1.796</td>
</tr>
<tr>
<td>Carnahan et al. (2016)</td>
<td>SC</td>
<td>US</td>
<td>J</td>
<td>N = 3</td>
<td>S</td>
<td>RG</td>
<td>MWR</td>
<td>3.8363</td>
</tr>
<tr>
<td>Cash (2013)</td>
<td>Q</td>
<td>US</td>
<td>D</td>
<td>N = 138</td>
<td>SS</td>
<td>PUB</td>
<td>MWR</td>
<td>0.501</td>
</tr>
<tr>
<td>Ciullo et al. (2015a)</td>
<td>SC</td>
<td>US</td>
<td>J</td>
<td>N = 4</td>
<td>S&amp;SS</td>
<td>RG</td>
<td>MWR</td>
<td>1.316</td>
</tr>
<tr>
<td>Ciullo et al. (2015b)</td>
<td>SC</td>
<td>US</td>
<td>J</td>
<td>N = 7</td>
<td>SS</td>
<td>RG</td>
<td>DNM</td>
<td>2.340</td>
</tr>
<tr>
<td>Guastello et al. (2000)</td>
<td>E</td>
<td>US</td>
<td>J</td>
<td>N = 124</td>
<td>S</td>
<td>RG</td>
<td>MWR</td>
<td>4.87*</td>
</tr>
<tr>
<td>Ledger (2003)</td>
<td>Q</td>
<td>US</td>
<td>D</td>
<td>N = 276</td>
<td>S</td>
<td>RG</td>
<td>MWR</td>
<td>0.328 .649</td>
</tr>
<tr>
<td>Study</td>
<td>Study Design</td>
<td>Study Location</td>
<td>Publication</td>
<td>N</td>
<td>Subject</td>
<td>Assessment Type</td>
<td>Study Quality</td>
<td>Hedges’ g ES</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------</td>
<td>----------------</td>
<td>-------------</td>
<td>-----</td>
<td>---------</td>
<td>-----------------</td>
<td>---------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Nair &amp; Narayanasamy  (2017)</td>
<td>Q</td>
<td>Malaysia</td>
<td>J</td>
<td>N = 70</td>
<td>SS</td>
<td>RG</td>
<td>MWR</td>
<td>2.703</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.485</td>
</tr>
<tr>
<td>Schenning et al.      (2013)</td>
<td>SC</td>
<td>US</td>
<td>J</td>
<td>N = 3</td>
<td>SS</td>
<td>RG</td>
<td>DNM</td>
<td>2.212</td>
</tr>
<tr>
<td>Tastan et al. (2008)</td>
<td>Q</td>
<td>Turkey</td>
<td>J</td>
<td>N = 50</td>
<td>S</td>
<td>RG</td>
<td>MWR</td>
<td>.458</td>
</tr>
<tr>
<td>Zakas et al. (2013)</td>
<td>SC</td>
<td>US</td>
<td>J</td>
<td>N = 3</td>
<td>SS</td>
<td>RG</td>
<td>MWR</td>
<td>.818</td>
</tr>
</tbody>
</table>

*Note. E = experimental, Q = quasi-experimental, SC = single case; J = journal, D = dissertation; S = Science, SS = Social Studies; RG = researcher generated, PUB = publisher created; M = meets, MWR = meets with reservation, DNM = does not meet

*This value has been winsorized.
<table>
<thead>
<tr>
<th>Study</th>
<th>Grade</th>
<th>Text Type</th>
<th>Text Level</th>
<th>TX Type</th>
<th>TX Inst</th>
<th>GO Type</th>
<th>Hedges’ g</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arslan (2006)</td>
<td>5</td>
<td>A: Course Reader</td>
<td>CD</td>
<td>GO Only</td>
<td>N: started intervention by working on concept maps prepared by the author</td>
<td>CM</td>
<td>0.115</td>
</tr>
<tr>
<td>Tastan et al. (2008)</td>
<td>11</td>
<td>RG: Conceptual Change</td>
<td>CD</td>
<td>GO Only</td>
<td>N: students taught how to create maps in previous unit; no teacher instruction during intervention</td>
<td>CM</td>
<td>.458</td>
</tr>
<tr>
<td>Berry et al. (2013)</td>
<td>3</td>
<td>A: Trade book</td>
<td>CD</td>
<td>MC: C</td>
<td>N: the creation of a class constructed concept map; no mention of teacher instruction</td>
<td>CM</td>
<td>0.919</td>
</tr>
<tr>
<td>Nair &amp; Narayanasamy (2017)</td>
<td>8</td>
<td>A: Textbook</td>
<td>CD</td>
<td>GO Only</td>
<td>Y: teacher taught students to complete CM in meaningful way</td>
<td>CM</td>
<td>2.703</td>
</tr>
<tr>
<td>Ledger (2003)</td>
<td>8</td>
<td>A: Textbook</td>
<td>8th grade</td>
<td>GO Only</td>
<td>N: students taught how to create CM prior to intervention</td>
<td>CM</td>
<td>0.328</td>
</tr>
</tbody>
</table>

### Mixed

<table>
<thead>
<tr>
<th>Study</th>
<th>Grade</th>
<th>Text Type</th>
<th>Text Level</th>
<th>TX Type</th>
<th>TX Inst</th>
<th>GO Type</th>
<th>Hedges’ g</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ledger (2003)</td>
<td>8</td>
<td>A: Textbook</td>
<td>8th grade</td>
<td>GO Only</td>
<td>N: students taught how to create CM prior to intervention</td>
<td>CM</td>
<td>0.649</td>
</tr>
<tr>
<td>Study</td>
<td>Char.</td>
<td>Grade</td>
<td>Text Type</td>
<td>Text Level</td>
<td>TX Type</td>
<td>TX Inst</td>
<td>GO Type</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------</td>
<td>-------</td>
<td>-----------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>Cash (2013)</td>
<td>GE,</td>
<td>6</td>
<td>A: Textbook</td>
<td>6th grade</td>
<td>MC: T, C, V</td>
<td>Y: models how to complete GO with real life examples and then models with text</td>
<td>MOD</td>
</tr>
<tr>
<td>Boon et al. (2006a)</td>
<td>GE,</td>
<td>10</td>
<td>A: Textbook</td>
<td>10th grade</td>
<td>MC: V, C</td>
<td>Y: teacher presentation w/ students filling in a GO</td>
<td>MOD</td>
</tr>
<tr>
<td>Boon et al. (2006b)</td>
<td>GE,</td>
<td>10</td>
<td>A: Textbook</td>
<td>10th grade</td>
<td>MC: V, C</td>
<td>Y: teacher presentation w/ students filling in a GO</td>
<td>MOD</td>
</tr>
<tr>
<td>LD, ID, or Low Achieving</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ciullo et al. (2015a)</td>
<td>LD,</td>
<td>4 &amp; 5</td>
<td>A: Basal Reader</td>
<td>Lexile 4th grade</td>
<td>MC: C, F, V</td>
<td>N: participants completed the missing parts of GO by locating information in the book</td>
<td>MOD</td>
</tr>
<tr>
<td>DiCecco &amp; Gleason (2002)</td>
<td>LD</td>
<td>6, 7, 8</td>
<td>A: Textbook</td>
<td>Middle School</td>
<td>MC: C, F, V</td>
<td>Y: students received direct instruction using GO showing explicit relationships for content material</td>
<td>MOD</td>
</tr>
<tr>
<td>Study</td>
<td>Char.</td>
<td>Grade</td>
<td>Text Type</td>
<td>Text Level</td>
<td>TX Type</td>
<td>TX Inst</td>
<td>GO Type</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-------</td>
<td>-------</td>
<td>-----------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Guastello et al. (2000)</td>
<td>LA</td>
<td>7</td>
<td>A: Textbook</td>
<td>CD</td>
<td>MC: C, TS</td>
<td>Y: instructor demonstrated how to organize info.; students created GO as they discussed with teacher</td>
<td>MOD</td>
</tr>
<tr>
<td><strong>Autism Spectrum Disorder</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zakas et al. (2013)</td>
<td>ASD &amp; ID</td>
<td>6 &amp; 8</td>
<td>RG: w/ Picture Cues</td>
<td>Lexile 3rd Grade</td>
<td>GO</td>
<td>N: used instructional scripts to guide students what to put on GO</td>
<td>MOD</td>
</tr>
<tr>
<td>Carnahan &amp; Williamson (2013)</td>
<td>ASD</td>
<td>7 &amp; 8</td>
<td>RG: 3 paragraphs</td>
<td>Controlled for passage length and reading level</td>
<td>MC: TS, V, C</td>
<td>N: overview of parts of Venn diagram</td>
<td>MOD</td>
</tr>
<tr>
<td>Schenning et al. (2013)</td>
<td>ASD &amp; ID</td>
<td>6 &amp; 7</td>
<td>RG: w/ Picture Cues</td>
<td>FK 2nd-3rd Grade</td>
<td>MC: C</td>
<td>Y: Instructor used explicit instruction to teach students inquiry process to complete GO</td>
<td>MOD</td>
</tr>
<tr>
<td>Carnahan et al. (2016)</td>
<td>ASD</td>
<td>Ages 16 &amp; 15</td>
<td>A: Science text &amp; trade book</td>
<td>6th grade Grade Level</td>
<td>MC: TS, C</td>
<td>N: discussed how GO could be used as thinking tool</td>
<td>MOD</td>
</tr>
</tbody>
</table>
Note: A = academic text that is not manipulated; RG = researcher generated text that was created by researcher; MC = multicomponent intervention that includes reading; V = vocabulary; C = comprehension, TS = Text Structure; CM = concept map; DISC = disciplinary graphic organizer; MOD = modal graphic organizer
Table 3

Meta-Regression Moderator Analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\beta$</th>
<th>SE</th>
<th>Z</th>
<th>$p$ value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study</td>
<td>.770</td>
<td>.706</td>
<td>1.09</td>
<td>.275</td>
<td>-613  2154</td>
</tr>
<tr>
<td>Grade</td>
<td>.224</td>
<td>.436</td>
<td>.51</td>
<td>.607</td>
<td>-631  1080</td>
</tr>
<tr>
<td>Instruction</td>
<td>1.445</td>
<td>.653</td>
<td>2.21</td>
<td>.027</td>
<td>.165  2725</td>
</tr>
<tr>
<td>Intervention Type</td>
<td>.0278</td>
<td>.673</td>
<td>0.04</td>
<td>.214</td>
<td>-1292 1348</td>
</tr>
</tbody>
</table>

Note: This model is based on $k = 16$ studies and $n = 18$ effect sizes; CI = confidence interval
PRISMA Chart of Study Search and Article Eligibility/Selection Procedures

**Identification**
- Records identified through database searching
  \( (n = 21,329) \)

**Screening**
- Records after duplicates removed
  \( (n = 15,681) \)
  - Abstracts excluded
    \( (n = 15,189) \)
    - Did not meet basic inclusion criteria \( (n = 14191) \)
    - Not published in English \( (n = 15) \)
    - Meta-Analysis or synthesis \( (n = 28) \)
    - Non-GO intervention \( (n = 271) \)
    - No Science or SS DV \( (n = 135) \)
    - Non-Experimental \( (n = 334) \)
    - Non-participants \( (n = 215) \)

**Eligibility**
- Abstracts screened (title and abstract)
  \( (n = 15,681) \)
- Full-text articles analyzed for eligibility
  \( (n = 492) \)
  - Full-text articles excluded
    \( (n = 476) \)
    - Participants \( (n = 76) \)
    - Non-experimental \( (n = 25) \)
    - Study design \( (n = 77) \)
    - Intervention features/IV \( (n = 161) \)
    - No measure of Science or Social Studies content knowledge \( (n = 132) \)
    - Meta-Analysis \( (n = 2) \)
    - Not Published in English \( (n = 2) \)
    - Data Included in Published Study \( (n = 1) \)

**Included**
- Articles included in review
  \( (n = 16) \)
- Electronic Search \( (n = 16) \)

Figure 2

Forest Plot of the Average Effect Size of Individual Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Effect Size with 95% CI</th>
<th>P-value</th>
<th>g</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arslan (2006)</td>
<td>0.12 [-0.30, 0.53]</td>
<td>0.586</td>
<td>.1150478</td>
</tr>
<tr>
<td>Ledger (2003)</td>
<td>0.25 [-0.01, 0.50]</td>
<td>0.063</td>
<td>.3282874</td>
</tr>
<tr>
<td>Tastan et al. (2008)</td>
<td>0.28 [0.05, 0.52]</td>
<td>0.018</td>
<td>.4575496</td>
</tr>
<tr>
<td>Cash (2013)</td>
<td>0.35 [0.16, 0.55]</td>
<td>0.000</td>
<td>.500824</td>
</tr>
<tr>
<td>Ledger (2003)</td>
<td>0.41 [0.24, 0.58]</td>
<td>0.000</td>
<td>.6491426</td>
</tr>
<tr>
<td>Zakas et al. (2013)</td>
<td>0.42 [0.25, 0.59]</td>
<td>0.000</td>
<td>.8184927</td>
</tr>
<tr>
<td>Berry et al. (2013)</td>
<td>0.47 [0.29, 0.66]</td>
<td>0.000</td>
<td>.9189973</td>
</tr>
<tr>
<td>Ciullo et al. (2015a)</td>
<td>0.49 [0.30, 0.69]</td>
<td>0.000</td>
<td>1.316401</td>
</tr>
<tr>
<td>DiCecco &amp; Gleason (2002)</td>
<td>0.56 [0.34, 0.78]</td>
<td>0.000</td>
<td>1.410394</td>
</tr>
<tr>
<td>Boon et al. (2006b)</td>
<td>0.73 [0.41, 1.06]</td>
<td>0.000</td>
<td>1.793659</td>
</tr>
<tr>
<td>Carnahan &amp; Williamson (2013)</td>
<td>0.77 [0.45, 1.10]</td>
<td>0.000</td>
<td>1.796239</td>
</tr>
<tr>
<td>Boon et al. (2006a)</td>
<td>0.90 [0.53, 1.27]</td>
<td>0.000</td>
<td>1.990411</td>
</tr>
<tr>
<td>Schenning et al. (2013)</td>
<td>0.96 [0.59, 1.33]</td>
<td>0.000</td>
<td>2.211692</td>
</tr>
<tr>
<td>Ciullo et al. (2015b)</td>
<td>1.08 [0.68, 1.48]</td>
<td>0.000</td>
<td>2.340093</td>
</tr>
<tr>
<td>Nair &amp; Narayanasamy (2017)</td>
<td>1.22 [0.78, 1.66]</td>
<td>0.000</td>
<td>2.703025</td>
</tr>
<tr>
<td>Nair &amp; Narayanasamy (2017)</td>
<td>1.38 [0.88, 1.88]</td>
<td>0.000</td>
<td>3.485499</td>
</tr>
<tr>
<td>Carnahan et al. (2016)</td>
<td>1.46 [0.96, 1.97]</td>
<td>0.000</td>
<td>3.836307</td>
</tr>
<tr>
<td>Guastello et al. (2000)</td>
<td>1.68 [1.07, 2.30]</td>
<td>0.000</td>
<td>4.87</td>
</tr>
</tbody>
</table>

Random-effects REML model
Figure 3

Funnel Plot for Publication Bias
Appendix A

Sample Graphic Organizers
Example Concept Map

- ice
  - is
  - like solid

- liquid
  - like forms

- gas
  - like forms

- Water
  - has made of elements
  - are hydrogen
  - are oxygen
Example Modal Graphic Organizer
Example Disciplinary Graphic Organizer

<table>
<thead>
<tr>
<th>People or Event</th>
<th>When or Where</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Action** What did this person do? What happened during this event?

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

**Continuity or Change** What stayed the same? What immediately changed as a result of this event or person’s actions?

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

**Historical Significance** What long term changed occurred as a result of this event or action? Why are we still talking about this specific event/person’s actions as part of history?

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>
Appendix B
Coding Manual
<table>
<thead>
<tr>
<th>Variable and Name for Data Entry</th>
<th>Description</th>
<th>Code/Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CODER IDENTIFICATION</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A. Coder</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CODER</td>
<td>• First and last initials of person completing the study coding</td>
<td>AG AC</td>
</tr>
<tr>
<td><strong>B. Code date</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CODE_DATE</td>
<td>• Date coding of the study was completed</td>
<td>02/13/12</td>
</tr>
</tbody>
</table>

**STUDY LEVEL VARIABLES**

The unit for coding consists of one study. One study is defined as a set of data obtained from a research sample or subsample of subjects who are compared with each other, as well as the treatments, measures, and statistical analyses applied to this data.

If one study is reported in different forms (e.g., same study reported in two different journal articles), coding should be done from all forms, with individual items coded using the form that provides the best information for that particular item or variable.

If a single report or article describes more than one study, each study should be coded separately and the publication year should be labeled accordingly with a numerical indicator (e.g., 2007.1, 2007.2).

**Study Identifiers**

<table>
<thead>
<tr>
<th><strong>C. Author(s)</strong></th>
<th>Description</th>
<th>Code/Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHOR</td>
<td>• List the author(s) of the study APA format, in the order they are listed in the publication.</td>
<td>Graham, S. &amp; Harris, K. Gersten, R., Baker, S., &amp; Edwards, L.</td>
</tr>
<tr>
<td><strong>D. Year of publication</strong></td>
<td>Year study was published, four digits. If there is more than one report for the same study, list the most recent publication year.</td>
<td>2011</td>
</tr>
<tr>
<td>YEAR_PUB</td>
<td>Multiple studies in one publication: 2007.1 2007.2</td>
<td></td>
</tr>
<tr>
<td><strong>E. Publication type</strong></td>
<td>The type of publication the study is presented in</td>
<td>J = journal article B = book chapter D = dissertation or thesis C = conference paper or presentation T = technical report O = Other (specify)</td>
</tr>
<tr>
<td><strong>F. Location of Study</strong></td>
<td>CD = cannot determine</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------------</td>
<td></td>
</tr>
<tr>
<td>STUDY_LOC</td>
<td>The country where the study was conducted</td>
<td></td>
</tr>
<tr>
<td>The United States</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Design, Assignment of participants, and location

<table>
<thead>
<tr>
<th><strong>G. Experimental design</strong></th>
<th>CD = cannot determine</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXP_DES</td>
<td>The type of experimental design used in the study</td>
</tr>
<tr>
<td>E = experimental, with random assignments of participants or classes to conditions (unit of analysis must be correct)</td>
<td></td>
</tr>
<tr>
<td>Q = quasi-experimental, with pretest data to indicate group equivalence</td>
<td></td>
</tr>
<tr>
<td>SS = single subject design</td>
<td></td>
</tr>
<tr>
<td>CD = cannot determine</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>H. SS design</strong></th>
<th>CD = cannot determine</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS_DES</td>
<td>Identify the type of SS study</td>
</tr>
<tr>
<td>T = treatment reversal/ABAB</td>
<td></td>
</tr>
<tr>
<td>M = multiple baseline</td>
<td></td>
</tr>
<tr>
<td>O = other</td>
<td></td>
</tr>
<tr>
<td>CD = cannot determine</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>I. Unit of assignment</strong></th>
<th>CD = cannot determine</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASSIGN</td>
<td>Identify the unit of assignment to conditions</td>
</tr>
<tr>
<td>IS = individual student</td>
<td></td>
</tr>
<tr>
<td>MP = matched pair</td>
<td></td>
</tr>
<tr>
<td>G = group (classrooms, schools, tutoring groups)</td>
<td></td>
</tr>
<tr>
<td>SD = school districts, regions</td>
<td></td>
</tr>
<tr>
<td>CD = cannot determine</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>J. Pretest equivalence</strong></th>
<th>CD = cannot determine</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRE_EQ</td>
<td>Was the equivalence of treatment and comparison group(s) prior to intervention reported on content knowledge?</td>
</tr>
<tr>
<td>Y = yes</td>
<td></td>
</tr>
<tr>
<td>N = no</td>
<td></td>
</tr>
<tr>
<td>CD = cannot determine (only permissible with experimental studies using random assignment or students as own control)</td>
<td></td>
</tr>
</tbody>
</table>

REMEMBER: There can be no more than ½ standard deviation difference between treatment and comparison conditions on a content pre-test. This is necessary for the study to be included.
| K. Locale LOC | • The community setting where the study was conducted | U = Urban  
S = Suburban  
R = Rural  
CD = Cannot determine |
| --- | --- | --- |

<table>
<thead>
<tr>
<th>SAMPLE/STUDENT CHARACTERISTICS</th>
</tr>
</thead>
</table>
| L. Grade level GRADE | • The grade level of students (1-12) in the study. If students are in more than one grade, list each grade separated by commas. | 6  
Multiple grades:  
6, 7, 8  
3, 7, 9 |
| M. Type of Student STUD_TYPE | • Specify the type of disability students have been identified with  
• For multiple populations coder will list student types by commas | LD = Learning disability  
ID = Intellectual disability  
ASD = Autism Spectrum Disorder  
ND = No disability  
BD = Behavior disability  
LA = Low Achieving (disability not identified)  
Multiple types: ASD, ID (please specify)  
CD = cannot determine |
| N. Identification of disability ID_DIS | • Specify how students were identified as having a disability (LD) | SI = school identified  
RI = researcher identified  
B = Both school and researcher identified  
ND = students are described as students with disabilities but there is no data to support this disability label  
CD = cannot determine |
| O. Pre-study fluency performance of students in treatment group REAF_TX | • Specify the fluency level of students prior to intervention | Y = data about students’ fluency is provided, specify tests and provide page number where data is reported  
CD = no data is provided regarding the writing performance of students |
| P. Pre-study fluency performance of students in treatment group | • Specify the fluency level of students prior to intervention | Y = data about students’ fluency is provided, specify tests and provide page number where data is reported  
CD = no data is provided regarding the writing performance of students |
| Q. Pre-study general comprehension performance of students in treatment group | • Specify the fluency level of students prior to intervention | Y = data about students’ fluency is provided, specify tests and provide page number where data is reported  
CD = no data is provided regarding the writing performance of students |
| R. Pre-study general comprehension performance of students in treatment group | • Specify the fluency level of students prior to intervention | Y = data about students’ fluency is provided, specify tests and provide page number where data is reported  
CD = no data is provided regarding the writing performance of students |

**INTERVENTION & CONTROL COMPONENTS**

| S. Content area | • Identify the content area in which writing was taught. | S = Science  
SS = History/Social Studies  
CD = Cannot determine |
| T. Subject | • Identify which subject within the larger discipline the GO was implemented | For Science:  
B = Biology  
C = Chemistry  
P = Physics  
A = Astronomy  
O = Other |
| U. Intervention Components | INT_COMP | Identify if the intervention only contained a graphic organizer or if it contained other reading components. | For Social Studies:  
H = History  
O = Other  
For All:  
CD = cannot determine  
GO = GO only reported  
MC = Multi-component (Graphic organizer and reading instruction/activities) |
|---------------------------|----------|-------------------------------------------------------------------------------------------------|---------------------------------------------------------------------|
| V. Type of non-GO reading instruction and/or practice for multicomponent interventions | INT_TYPE | If the intervention is multi-component, identify the other type of reading instruction/practice provided to students | V = Vocabulary  
F = Fluency  
C = Comprehension  
O = other, specify  
CD = cannot determine |
| W. Treatment location | TX_PLACE | Specifies the type of learning environment that intervention and comparison conditions were conducted in.  
Note: If treatment and control were delivered in different types of environments, please note on the coding sheet | GE = General education classroom  
SP = Special education classroom  
RP = Resource room/pull-out from regular classroom  
AS = After school program  
L = Laboratory  
O = Other, specify  
CD = Cannot determine |
| X. Delivery of the Intervention | DELIVER | Specify who delivered the intervention to the students in the treatment and control conditions  
Note: If treatment and control were delivered in different types of environments, please note on the coding sheet | CT = content area teacher  
LT = language arts teacher  
R = researcher/research assistant  
O = Other, specify  
CD = cannot determine |
| Y. Training for implementers | TRAIN | Were those individuals who delivered the intervention(s) provided training to do so? This must be explicitly stated in the text. (e.g., Teachers received 3 hours | Y = yes  
N = no  
CD = cannot determine |
| Z. Length of training for implementers | Length of the training or professional development for implementers (specify unit, e.g., days, weeks, hours) | 6 days 3.5 hours
| CD = cannot determine |
| AA. Treatment group | Short description (a maximum of 3 sentences) to explain what students did in the treatment condition. | Students received graphic organizer and vocabulary instruction on signal phrases during the intervention. |
| BB. Comparison group(s) | Short description of comparison condition or baseline, including descriptions for what students did in no-treatment control conditions, if this information is provided. | Students participated in a no-treatment control condition that involved reading the text without any support; The control condition received a general notetaking intervention. |
| CC. Baseline instruction | Short description of any reading or graphic organizer instruction students received prior to the treatment or in the baseline phase. | Key terms and grammar terms were previewed using pictures and maps during the baseline phase. |
| DD. Intervention instruction other | Short description of any other content area instruction reported during the intervention | Students watched a 10 minute video. |

**INSTRUCTIONAL COMPONENTS**

| EE. Teacher Instruction | Did the instructor provide instruction before students worked on the graphic organizers in small groups or independently? Examples include: think-alouds, explicit modeling of the strategy etc. |
| Did the teacher presented the graphic organizer to the students |
| N= it was not reported that the teacher presented the graphic organizer to the students before group or individual practice |
| FF. Supported Practice | Did the students practice with graphic organizers with instructor support? | Y= students practiced with the graphic organizer with instructor support  
N= practice with the graphic organizer with instructor support was not reported |
|-----------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------|
| GG. Collaborative Activity | Did the students engage in a collaborative activity with the graphic organizers? This could be co-creating graphic organizers or studying graphic organizers together. | Y= students engaged in collaborative practice  
N= collaborative practice was not reported |
| HH. Individual Practice | Did the students engage in individual practice with the graphic organizers? | Y= students engaged in individual practice  
N= individual practice was not reported |

### INTERVENTION INTENSITY & FIDELITY

| II. Size of Intervention Group | Number of students in intervention group | 3 students  
CD = cannot determine |
|-------------------------------|----------------------------------------|--------------------------------------------------|
| JJ. Duration of intervention | The total length of the intervention, not including days used for pre- and post-testing. (specify unit, e.g., days, weeks, hours) | 15 days  
1 month  
CD = cannot determine |
| KK. Frequency of intervention | The number of times the intervention occurred (this is usually per week, but specify unit, e.g., per day, per week, per month) | 2 times per week  
1 time per day  
3 times per month  
CD = cannot determine |
| LL. Dosage of the intervention | The length of each intervention session (specify unit, e.g., minutes, hours) | 30 minutes  
2 hours  
CD = cannot determine |
| MM. Implementation fidelity | Was the fidelity of treatment implementation reported? | YA = was reported, above 0.80  
YB = was reported, below 0.80 |
| NN. Graphic Organizer Type | GO_TYPE | • What type of graphic organizer or graphic organizers were used during the intervention? | M = Modal (organized by pattern of thinking or text structure (e.g. problem/solution, compare/contrast))  
CM = Concept map (key concepts are identified and lines are drawn to show relationships between concepts; words are placed on lines to link words)  
DIS = information organized by key concepts, theories, or themes (e.g. historical significance)  
CD = cannot determine |
|---------------------------|---------|--------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| OO. Knowledge Representation | GO_ORG | • If Modal GO, Identify how information was connected within the graphic organizer | H = Information connected by relationship, hierarchy  
C/C = Information connected by relationship, compare & contrast  
C/E = Information connected by relationship, cause and effect  
P = Information connected by relationship, process  
P/S = information connected by problem/solution  
CD = cannot determine |
| PP. Picture of Graphic Organizer | | • Was a picture of the graphic organizer provided in the study? | Y = yes  
N = No |
| GO_PIC |
|---|---|---|
| **QQ. Graphic Organizer Task** | • Identify how the student graphic organizer(s) was completed. | SCR= Construct from scratch  
TEMP= template that students filled in  
PC= partially completed; a template was provided and some information on the graphic organizer was already filled in |
| **RR. Mapper** | • Identify who completed the student graphic organizer. | STU = Student  
TEA= Teacher  
REA = Researcher |
| **SS. Type of Response** | • Identify the type of response used to complete the graphic organizer.  
If more than one mode of response was used, specify both. | PEN = Pencil & paper  
COM = Computer  
ORA = Oral |

| TEXT |
|---|---|---|
| **TT. Text Description** | • Provide a brief description of the text or texts used | Example:  
Getting to Know Our Body  
Chapter in course book  
"Science Curriculum for Elementary Students" |
| **UU. Type of Text** | • Identify the type of text used in the intervention | TEXT = Textbook  
TRADE = Tradebook  
BR = Basal reader  
RG= researcher generated  
O = Other, specify  
CD = cannot determine |
| **VV. Grade** | • Identify how the text is leveled and the level of the text (e.g. grade, Lexile) | Examples:  
2\textsuperscript{nd}-3\textsuperscript{rd} grade  
8\textsuperscript{th} grade  
CD= cannot determine |
## WW. Text Modification

**TEXT_MOD**

- Identify if the text is an original academic text or has been modified or created by the researcher.

| O | Original academic text |
| N | No modification |
| RM | Researcher Modified (text modified or reduced to lower fluency level) |
| PC | Picture cues added |
| RC | Researcher Created |
| CD | Cannot determine |

## XX. Text Reader

**TEXT_READ**

- Identify who read the text as students completed the student graphic organizers.

| T | Teacher read text |
| S | Student read text |
| O | Other student read text |
| R | Researcher read text |
| CD | Cannot determine |

## DEPENDENT VARIABLES

Create one coding sheet for each dependent measure of *content learning* that is reported. For example, if a study reports data for a written measure of content learning and data for a multiple choice content learning outcome measure, 2 coding lines are necessary. Additionally, the 4 variables following Measure Name (type of measure, reliability, time of measurement, and time after the intervention) should be completed for each measure of content learning and denoted with the appropriate number.

**NOTE:** General comprehension measures and written measures that also measure students' ability to compose are excluded.

### YY. Measure Name

**MEAS_NAME**

- The name of the content measure and a brief, 15 words or less, description of the measure.

| Measure: 10 MC question assessment of content learning created by researcher |

### ZZ. Type of measure

**QUES_TYPE**

- The type of questions & format of measure

| MC | Multiple choice |
| SA | Short answer |
| WM | Essay or summary |
| INT | Interview |
| MT | Maze Task |
| O | Please specify |
| CD | Cannot determine |

### AAA. Measure Created by

**MEAS_TYPE**

- Type of Measure

<p>| ST | Standardized measure |
| RG | Researcher generated |</p>
<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBB. Student Response</td>
<td>How did the student provide answers on the assessment?</td>
<td>W=written, O=oral, CD=Cannot determine</td>
</tr>
<tr>
<td>CCC. Student Support with Measure</td>
<td>Please describe any support students received with the measure.</td>
<td>Example: Student allowed to look back in text if incorrect responses provided</td>
</tr>
<tr>
<td>DDD. Reliability of measure</td>
<td>Reliability of the outcome measure (split-half, inter-rater, etc.) reported in decimal form. If multiple reliabilities are reported, record all of them, denoting each by which type of reliability it is.</td>
<td>One type of reliability reported: 0.92 More than one type reported: 0.88 split-half 0.94 inter-rater CD = cannot determine, not reported (This is blank in Stata)</td>
</tr>
<tr>
<td>EEE. Timing of measure</td>
<td>Specify when the assessment occurred</td>
<td>PO = posttest only PP = pre and posttest M = maintenance G = generalization or transfer CD = cannot determine</td>
</tr>
<tr>
<td>FFF. Timing of measure after intervention</td>
<td>How long was the assessment given after the end of the intervention? (specify unit, e.g., days, weeks, hours)</td>
<td>1 week 3 days 2.5 months posttest given immediately after intervention = 0</td>
</tr>
</tbody>
</table>
Studies with more than two conditions or comparisons will need one effect size coding sheet for each comparison made. For example, a study that presents data comparing one writing intervention, a different writing intervention, and a no-treatment control would require three effect size coding sheets. One effect size coding sheet would be used to compare treatment 1 and the control, one would be used to compare treatment 2 and control, and one would be used to compare treatment 1 and treatment 2. Additionally, if a study contains multiple measures, one effect size sheet is required for each measure. Using the example above, if the study contained two measures of writing quality, it would require 6 effect size sheets. Remember, for each effect size sheet, you are only coding variables for the two measures compared on that sheet, not for the entire study (if there are other measures or other comparisons). Round all non-whole numbers to three decimal places.

<table>
<thead>
<tr>
<th>GGG. Participants assigned to treatment condition</th>
<th>Specify the number of participants assigned to the treatment condition.</th>
<th>12 145</th>
</tr>
</thead>
<tbody>
<tr>
<td>N_ASSIGN_TX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHH. Participants assigned to comparison condition</td>
<td>Specify the number of participants assigned to the comparison condition.</td>
<td>12 145</td>
</tr>
<tr>
<td>N_ASSIGN_COM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III. Participants observed in the treatment condition</td>
<td>Specify the number of participants in the treatment condition who were observed/scored post-intervention on the writing measure being coded</td>
<td>10 139</td>
</tr>
<tr>
<td>N_OBS_TX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JJJ. Participants observed in the comparison condition</td>
<td>Specify the number of participants in the comparison condition who were observed/scored post-intervention on the writing measure being coded</td>
<td>10 139</td>
</tr>
<tr>
<td>N_OBS_COM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KKK. Total observed sample</td>
<td>Specify the total observed sample size in the study (add participants observed in treatment and participants observed in control)</td>
<td>20 278</td>
</tr>
<tr>
<td>N_TOTAL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLL. Pre-test mean for treatment group</td>
<td>Specify the pre-test mean for the treatment group (leave blank if there was no pretest).</td>
<td>7.5682 = 7.568 100</td>
</tr>
<tr>
<td>PRE_M_TX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMM. Pre-test standard deviation for treatment group</td>
<td>Specify the pre-test standard deviation for the treatment group (leave blank if there was no pretest).</td>
<td>4.5799 = 4.5806</td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>NNN. Post-test mean for treatment group</td>
<td>Specify the post-test mean for the treatment group.</td>
<td>10.67115</td>
</tr>
<tr>
<td>POST_M_TX</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>OOO. Post-test standard deviation for treatment group</td>
<td>Specify the post-test standard deviation for the treatment group.</td>
<td>2.31</td>
</tr>
<tr>
<td>POST_SD_TX</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>PPP. Pre-test mean for comparison group</td>
<td>Specify the pre-test mean for the comparison group (leave blank if there was no pretest).</td>
<td>7.5449 = 7.545110</td>
</tr>
<tr>
<td>PRE_M_COM</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>QQQ. Pre-test standard deviation for comparison group</td>
<td>Specify the pre-test standard deviation for the comparison group (leave blank if there was no pretest).</td>
<td>22.89</td>
</tr>
<tr>
<td>PRE_SD_COM</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>RRR. Post-test mean for comparison group</td>
<td>Specify the post-test mean for the comparison group.</td>
<td>8.72111</td>
</tr>
<tr>
<td>POST_M_COM</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>SSS. Post-test standard deviation for comparison group</td>
<td>Specify the post-test standard deviation for the comparison group.</td>
<td>2.675</td>
</tr>
<tr>
<td>POST_SD_COM</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>TTT. Type of effect size calculated</td>
<td>The type of effect size calculated SMD = standardized mean difference (post-test data only) GS = gain score (pre- and post-test data) BCSMD= between-case standardized mean difference estimator O = other, specify</td>
<td>0.781</td>
</tr>
<tr>
<td>ES_TYPE</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>UUU. Effect Size</td>
<td>Effect size value, rounded to three decimal places reported as d</td>
<td></td>
</tr>
<tr>
<td>ES</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
</tbody>
</table>
### EFFECT SIZES – SS

<table>
<thead>
<tr>
<th>VVV. Type of effect size calculated</th>
<th>The type of effect size calculated</th>
<th>SMD = standardized mean difference (post-test data only) GS = gain score (pre- and post-test data) O = other, specify</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES_TYPE</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>XXX. Effect Size</th>
<th>Effect size value, rounded to three decimal places</th>
<th>0.781</th>
</tr>
</thead>
</table>

*For single subject studies, please copy output into corresponding columns for SS from the SCDHLM app.

### Group Design Study Quality

<table>
<thead>
<tr>
<th>YYY. Random Assignment</th>
<th>Identify if participants are assigned to treatment through a random process?</th>
<th>Y = RCT or random assignment N = Quasi-experimental CD = Cannot determine</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAND_ASSIGN</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### SAMPLE ATTRITION

<table>
<thead>
<tr>
<th>ZZZ. Attrition</th>
<th>What percentage of the original participants in treatment group completed the study?</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATTRITION_TX</td>
<td></td>
<td>CD = cannot determine, attrition data not reported</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AAAA. Attrition</th>
<th>What percentage of the original participants in comparison group completed the study?</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATTRITION_COM</td>
<td></td>
<td>CD = cannot determine, attrition data not reported</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BBBB. Equal Attrition</th>
<th>Was attrition equal across treatment and comparison(s) groups? Equal = no more than 10% difference in attrition between groups</th>
<th>Y = yes N = no CD = cannot determine, not reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATT_EQ</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**See Pretest equivalence PRE_EQ column for additional QI information.
<table>
<thead>
<tr>
<th>Single Subject Study Quality</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CCCC. Independent Variable</strong></td>
<td><strong>IV_MAN</strong>&lt;br&gt;• Is the independent variable systematically manipulated, with the researcher determining when and how the IV conditions change?</td>
</tr>
<tr>
<td><strong>DDDD. Inter-assessor agreement</strong></td>
<td><strong>IOA</strong>&lt;br&gt;• Is IOA assessed in each condition and in each phase of the design?</td>
</tr>
<tr>
<td><strong>EEEE. Inter-assessor agreement_raters</strong></td>
<td><strong>IOA_RATERS</strong>&lt;br&gt;• How many assessors/observers measure the behaviors of interest/outcomes?&lt;br&gt;• If different numbers of raters assessed at each phase of the design, report accordingly (see example 2 to the right).</td>
</tr>
<tr>
<td><strong>FFFF. Inter-assessor agreement statistics</strong></td>
<td><strong>IOA_STATS</strong>&lt;br&gt;• Specify all IOA reported (commonly reported in percent agreement or Cohen’s kappa).&lt;br&gt;• Distinguish statistics by phase or condition of study</td>
</tr>
<tr>
<td><strong>GGGG. Visual data display</strong></td>
<td><strong>DATA_DISPLAY</strong>&lt;br&gt;• Does the study include a visual display of participants’ data points for each phase of the design?</td>
</tr>
<tr>
<td><strong>HHHH. Number of phases</strong></td>
<td><strong>SS_PHASES</strong>&lt;br&gt;• How many phases were there in the study design?&lt;br&gt;• Provide a brief description of the phases.</td>
</tr>
<tr>
<td><strong>IIII. Number of data points collected per phase</strong></td>
<td><strong>SS_DATAPTS</strong>&lt;br&gt;• How many data points were collected at each phase?</td>
</tr>
</tbody>
</table>
REFERENCES


https://doi.org/10.1598/RT.64.5.9


https://doi.org/10.1177/1053451213480025


https://doi.org/10.1177/016264340602100101


https://doi.org/10.1080/00461520.2014.916217

https://doi.org/10.1080/09362835.2013.802236


[https://doi.org/10.1177/0014402919893932](https://doi.org/10.1177/0014402919893932)

[https://doi.org/10.1177/0145445514552890](https://doi.org/10.1177/0145445514552890)

[https://doi.org/10.1177/0731948713516767](https://doi.org/10.1177/0731948713516767)

[https://doi.org/10.1177/0022219414539566](https://doi.org/10.1177/0022219414539566)

[https://doi.org/10.1111/ldr.12017](https://doi.org/10.1111/ldr.12017)


[https://doi.org/10.1080/02702710802274903](https://doi.org/10.1080/02702710802274903)

Darch, C. B., Carnine, D. W., & Kameenui, E. J. (1986). The role of graphic organizers and social

https://doi.org/10.1177/002246698602000305

https://doi.org/10.2307/1510469

https://doi.org/10.1177/073194871103400104

https://doi.org/10.1111/j.1540-5826.2011.00341.x

https://doi.org/10.1177/00222194020350040201


learners who struggle. *Topics in Language Disorders, 32*(1), 69-84.

[https://doi.org/10.1097/tld.0b013e318245618e](https://doi.org/10.1097/tld.0b013e318245618e)


[https://doi.org/10.1177/00222194070400030301](https://doi.org/10.1177/00222194070400030301)


[https://doi.org/10.1177/0014402914527238](https://doi.org/10.1177/0014402914527238)


[https://doi.org/10.1177/001440290607200301](https://doi.org/10.1177/001440290607200301)


[https://doi.org/10.1177/001440290507100202](https://doi.org/10.1177/001440290507100202)


[https://doi.org/10.1177/074193250002100605](https://doi.org/10.1177/074193250002100605)


Historical Thinking Project, (n.d.). *Historical thinking concepts*. Historical Thinking Project.


https://doi.org/10.1037/edu0000082

https://historicalthinking.ca/historical-thinking-concepts

https://doi.org/10.1177/001440290507100203

https://doi.org/10.1111/lrdp.12025

https://doi.org/10.1177/00222194040370020201

https://doi.org/10.1007/s10803-014-2280-x


DOI: 10.1093/oxfordhb/9780199734689.013.0004


https://doi.org/10.1111/j.1540-5826.2011.00342.x


https://doi.org/10.1111/ldrp.12026


https://doi.org/10.1136/bmj.b2535


https://doi.org/10.3102/00346543076003413


https://doi.org/10.1177/016264349601300204


https://doi.org/10.1002/sce.1029
Pustejovsky, J. E. (2016). scdhlm: A web-based calculator for between-case standardized mean differences (Version 0.3.1) [Web application].

https://jepusto.shinyapps.io/scdhlm


https://doi.org/10.3102/1076998614547577


https://doi.org/10.1080/09362835.2016.1196441


https://doi.org/10.1177/0272431689091009


https://doi.org/10.1111/j.1540-5826.2008.00264.x


https://doi.org/10.1016/j.rasd.2012.12.007
https://doi.org/10.1007/S10648-017-9403-9

https://doi.org/10.1177/0741932508327465

https://doi.org/10.17763/haer.78.1.v62444321p602101


https://doi.org/10.1177/0741932518770288

https://doi.org/10.1177/0022219412451131

https://doi.org/10.1002/jrsm.1091

**Tastan, I., Dikmenli, M., & Cardak, O. (2008). Effectiveness of the conceptual change texts accompanied by concept maps about students' understanding of the molecules**


https://doi.org/10.4073/cmdp.2016.1


https://doi.org/10.1177/0731948718806654


https://doi.org/10.1007/s11145-017-9802-9
https://doi.org/10.1037/0022-0663.97.4.538


https://doi.org/10.1177/1540796915592155

https://doi.org/10.1016/j.rasd.2013.06.001

https://doi.org/10.3109/17518423.2015.1100690