The dissertation of Chelsea E. Cameron was reviewed and approved* by the following:

Peggy N. Van Meter  
Associate Professor of Education  
Dissertation Adviser  
Chair of Committee

Rayne A. Sperling  
Professor of Education

Crystal M. Ramsay  
Research Project Manager  
Affiliate Assistant Professor of Educational Psychology

Michael J. Milligan  
Teaching Professor of History

Pui-Wa Lei  
Professor in Charge, Educational Psychology

*Signatures are on file in the Graduate School.
Abstract

College students often struggle to learn from multiple documents without instructional support. Previous research suggests that self-questioning strategies can aid comprehension and that questions aimed at multiple documents can support multiple document learning. Multiple Document Question Instruction (MDQI) was developed to teach students how to generate high-quality questions to support multiple document learning. A prior study tested the effects of this instruction, in which experimental participants received MDQI after reading the texts. Although positive effects of instruction were found, effects were small. The current study replicates that study, but includes a third condition in which participants receive MDQI before reading the documents. Findings suggest MDQI supports the learner’s abilities to generate high-quality questions and that the generation of such questions, apart from condition, does have a moderate effect on multiple document learning. While there were significant effects of condition on question generation, there were no significant condition effects on multiple document learning. There was an important pattern, however, in that MDQI before reading performed worse than both MDQI after reading and control conditions on a measure of multiple document integration and learning, contrary to expectations. Implications for further development of MDQI, educational use, and future research are discussed.
# TABLE OF CONTENTS

List of Tables........................................................................................................................................ ix

List of Figures........................................................................................................................................ x

Chapter 1. INTRODUCTION........................................................................................................................ 1

Statement of Purpose.................................................................................................................................. 1

The Need to Support Multiple Document Learning.................................................................................... 2

Researcher-Generated Tasks Versus Learner-Directed Strategies.............................................................. 3

Purpose of the Current Research.............................................................................................................. 6

Chapter 2. LITERATURE REVIEW.............................................................................................................. 11

Introduction................................................................................................................................................ 11

Theoretical Framework.............................................................................................................................. 11

MD-TRACE................................................................................................................................................. 11

Theoretical principles and assumptions..................................................................................................... 12

The Construction-Integration Model.......................................................................................................... 12

The Documents Model............................................................................................................................... 13

Elements of reader processing.................................................................................................................... 13

Internal factors.......................................................................................................................................... 13

External factors.......................................................................................................................................... 14

Research on Learning from Multiple Documents...................................................................................... 14

Defining a document.................................................................................................................................. 15

Measurement of Multiple Documents Learning.......................................................................................... 16

Process measures of multiple documents learning.................................................................................... 17
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Think-aloud protocols</td>
<td>17</td>
</tr>
<tr>
<td>Product measures of multiple document learning</td>
<td>17</td>
</tr>
<tr>
<td>Verification tasks</td>
<td>18</td>
</tr>
<tr>
<td>Essays</td>
<td>19</td>
</tr>
<tr>
<td>Conclusions</td>
<td>20</td>
</tr>
<tr>
<td>Supporting Multiple Document Learning</td>
<td>21</td>
</tr>
<tr>
<td>Strategy use and multiple document learning</td>
<td>22</td>
</tr>
<tr>
<td>Question Generation and Multiple Document Learning</td>
<td>24</td>
</tr>
<tr>
<td>The Current Study</td>
<td>26</td>
</tr>
<tr>
<td>Chapter 3. METHODS</td>
<td>30</td>
</tr>
<tr>
<td>Participants</td>
<td>30</td>
</tr>
<tr>
<td>Design</td>
<td>30</td>
</tr>
<tr>
<td>Materials</td>
<td>31</td>
</tr>
<tr>
<td>Individual Difference Measures</td>
<td>31</td>
</tr>
<tr>
<td>Demographic survey</td>
<td>31</td>
</tr>
<tr>
<td>Reading comprehension</td>
<td>31</td>
</tr>
<tr>
<td>Prior knowledge</td>
<td>32</td>
</tr>
<tr>
<td>Document Knowledge Posttest Measure</td>
<td>32</td>
</tr>
<tr>
<td>Instructional Materials</td>
<td>34</td>
</tr>
<tr>
<td>Documents</td>
<td>34</td>
</tr>
<tr>
<td>Documents matrix</td>
<td>36</td>
</tr>
<tr>
<td>Experimental MDQI materials</td>
<td>37</td>
</tr>
<tr>
<td>Instructional video</td>
<td>37</td>
</tr>
</tbody>
</table>
Question frameworks................................................................................................. 38
Control materials........................................................................................................ 38
Question generation measure..................................................................................... 39
Question selection measure......................................................................................... 39
Document integration essay measure.......................................................................... 39
Procedure........................................................................................................................ 40
Question coding rubric.................................................................................................. 42
Document integration essay coding rubric...................................................................... 45
Chapter 4. RESULTS...................................................................................................... 49
Preliminary Analyses....................................................................................................... 49
Question Generation....................................................................................................... 50
Question Selection........................................................................................................... 52
Condition Effects on Learning....................................................................................... 53
MDQ Generation Effects on Learning........................................................................... 56
Summary of Results......................................................................................................... 59
Chapter 5. DISCUSSION................................................................................................ 61
Summary of Findings....................................................................................................... 61
Question Generation....................................................................................................... 61
Question Selection........................................................................................................... 62
Condition Effects on Learning....................................................................................... 62
MDQ Generation Effects on Learning........................................................................... 63
Conclusions..................................................................................................................... 63
Implications of Findings................................................................................................. 65
Self-directed multiple document learning.................................................. 65
Questions as tools for learning................................................................. 65
Timing matters......................................................................................... 66
Limitations.............................................................................................. 66
Future Research...................................................................................... 69
Conclusion.............................................................................................. 70
References............................................................................................. 71
Appendix A: Study Materials & Measures............................................... 78
  Implied Informed Consent..................................................................... 79
  Demographics Survey........................................................................... 80
  Reading Comprehension Measure....................................................... 82
  Prior Knowledge Measure.................................................................. 88
  Question Generation & Selection Directions...................................... 91
  Document Integration Measure: Essay Response Prompt.................. 92
  Document Knowledge Posttest......................................................... 93
Appendix B: Experimental Documents.................................................... 96
  Document 1: William Jennings Bryan Opposes U.S. Occupation of the
    Philippines, 1990............................................................................ 97
  Document 2: Albert Beveridge Defends U.S. Imperialism, 1900............ 97
  Document 3: Photograph 1 American Soldiers in the Philippines........ 100
  Document 4: Political Cartoon 1 United States Imperialism............... 100
  Document 5: Political Cartoon 2 Showing Light to the Filipino People... 101
  Document 6: Photograph 2 Filipino Nationalists Work Toward Independence... 101
Appendix C: Multiple Document Question Instruction Materials................................. 102

Introduction................................................................................................................... 103

Video Transcript....................................................................................................... 103

Example Question Frameworks............................................................................... 105
List of Tables

Table 1 Description of Experimental Documents.......................................................... 33
Table 2 Order of Study Activities by Condition............................................................... 35
Table 3 Question Coding Rubric.................................................................................... 41
Table 4 Response Coding Rubric.................................................................................. 44
Table 5 Means and Standard Deviations for the Prior Knowledge, Reading Comprehension, & Questions................................................................. 47
Table 6 Question Type Frequencies, Proportions of Question Types, and Standardized Residuals................................................................................................ 52
Table 7 Question Generation and Selection Frequencies, Proportions, and Standardized Residuals......................................................................................... 53
Table 8 Response Type Frequencies, Proportions, and Standardized Residuals by Condition........................................................................................................... 55
Table 9 Condition Means and Standard Deviations for the Document Knowledge Posttest..................................................................................................................... 56
Table 10 MDQ Response Type Frequencies, Proportions, and Standardized Residuals.... 58
Table 11 MDQ Means and Standard Deviations for the Document Knowledge Posttest......................................................................................................................... 59
List of Figures

Figure 1 Documents Matrix........................................................................................................ 37
Chapter 1

Introduction

Statement of Purpose

The ability to critically consume information from multiple sources has become a crucial skill for the modern learner (Goldman & Scardamalia, 2013). The magnitude of information that is available within a simple web-browser search alone depicts the value of this ability. Unfortunately, college students struggle to select and integrate information from multiple documents to understand complex issues (e.g., Cerdán & Vidal-Abarca, 2008; Wineburg, 1991). Because this ability is necessary for multiple document learning and informed decision making, preparations for 21st-century reading and learning call for the modern learner to be able to read and consume information strategically, to become life-long learners and critical consumers of knowledge (Alexander, 2012).

While the ability to digest the massive amount of available information is a daunting task, there is great value in using and learning from multiple documents rather than from a single document alone (Nokes, Dole, & Hacker, 2007; Perfetti, Rouet, & Britt, 1999). Multiple documents can offer the learner the chance to explore the complexities of a topic by providing both complementary and conflicting information. In order to fully benefit from this information, however, the learner must engage in internal cognitive processes such as selecting relevant information and organizing this intertextual information into an integrated understanding of the documents.

Although previous research has often focused on the role of external factors on learners’ ability to integrate information from multiple documents, internal cognitive processes must also be fostered and developed through instructional supports such as strategy instruction. The current
research tests the effects of a question-generation strategy instruction intervention aimed at fostering these internal processes to support multiple document learning.

**The Need to Support Multiple Document Learning**

Although the modern learner frequently engages with multiple documents in both formal and informal settings, researchers have found that students often struggle to develop an integrated understanding of the information presented across sources (Brit & Sommer, 2004; Wiley & Voss, 1999). This level of multiple document learning requires the coordination of strategic processing to engage in the cognitive processes of selecting relevant document information, organizing document information, and integrating that information across documents (Bråten, Anmarkrud, Brandmo, & Strømsø, 2014).

To understand the full range of internal and external factors involved with multiple document learning, Rouet and Britt (2011) proposed the Multiple-Documents Task Relevancy and Content Extraction (MD-TRACE) model. The MD-TRACE model supposes that when a learner engages with multiple documents, external factors (e.g., the task specifications) and internal factors (e.g., strategy knowledge) impact the learner’s reading process. Previous studies have provided empirical support demonstrating the role of the task in multiple document learning such as provided questions (e.g., Britt & Sommer, 2004) and writing prompts (e.g., Gil, Bråten, Vidal-Abarca, & Strømsø, 2010a). There are other studies that show how internal factors, such as strategy use, impact multiple document learning (Bråten & Strømsø, 2011; Hagen, Braasch, & Bråten, 2014). In a study by Hagen, Braasch, and Bråten (2014), undergraduate students studied seven climate change texts that presented conflicting information on the topic. Participants took notes while reading the documents and then completed a summary or argument writing task. Findings indicated that the use of multiple text strategies such as elaboration during notetaking
and cross-text relationship awareness had positive associations with multiple text comprehension. External and internal factors also interact during the multiple document learning process and this interaction can impact how a learner engages with the materials. Bråten and Strømsø (2003), for example, found that law students’ depth of processing, time on task, and strategy use varied as knowledge grew and as the times of the exam approached while completing course readings. The Bråten and Strømsø (2003) study provides evidence of the internal and external factors interaction in that learners shift their reading behaviors based on the task.

**Researcher-Generated Tasks Versus Learner-Directed Strategies**

There is theoretical and empirical evidence that multiple document processing can be improved by providing supportive tasks such as question prompts (e.g., Cerdán & Vidal-Abarca, 2008) and argument or summary essay prompts (i.e., Gil et al., 2010b; Le Bigot & Rouet, 2007). Although this body of research has shown some promise, two significant weaknesses decrease confidence that task manipulations alone can address learners' needs when it comes to learning from multiple documents. First, the research has not only failed to identify a single task that is consistently effective, but also contains inconsistencies in task effectiveness across studies. In a study by Gil et al. (2010a), for example, undergraduate students studied five documents on climate change that presented conflicting information and completed a 21-item prior knowledge assessment. After studying the documents and completing the prior knowledge assessment, participants completed a writing task from either a summary essay prompt or argumentation essay prompt. Finally, participants completed the sentence verification task measures, these are measures of multiple text comprehension that require the participants to identify sentences or relationships that exist in the texts they read. Findings indicate that there is a significant
interaction between task and prior knowledge. Argument task participants with high prior knowledge performed better on the sentence verification measures and demonstrated more integration and deeper processing of the multiple texts in the written response than high prior knowledge summary task participants. But, summary task participants with low prior knowledge outperformed the low prior knowledge argument task participants on the written response. There were no differences between low prior knowledge task conditions on the sentence verification task measures. In another study by Gil et al. (2010b), undergraduate students studied seven documents on climate change that presented conflicting information. After studying the documents, participants completed a writing task from either a summary prompt or argumentation prompt, and then completed the sentence verification task measures. It was expected that the argumentation essay task would facilitate a deeper and more integrated text understanding than the summary essay task. The summary task participants, however, outperformed the argumentation task participants on sentence verification task measures. Summary task participants also included a greater number of sources and transformation of information than argument task participants, measures of integration and depth of processing of the multiple texts. These studies demonstrate that task alone cannot provide a singular answer to how to improve learning from multiple documents. Second, and more importantly, reliance on tasks as the driver of learners' multiple document learning prohibits development of learners' self-directed strategies for multiple document learning. In this respect, learners may be successful when external forces, such as a teacher, assign a particular task but they are unable to manage multiple documents without that additional support. Of course, such self-directed learning requires that the learner has the knowledge and ability to execute strategies for multiple document reading.
If the goal is to equip learners with the skills necessary to learn from multiple documents, it is necessary to explore the effects of learner-directed strategies on multiple document learning. Both the knowledge and regulation of strategies are essential components to the ideal processing of multiple documents. The MD-TRACE (Britt & Rouet, 2011) model states that multiple document reading is not only impacted by external factors such as tasks but also students' independent comprehension abilities. The development of independent comprehension abilities requires instruction to help learners read intertextually (Van Meter & Firetto, 2008) and development of deep-level intertextual strategies (Bråten et al., 2014). Forrest-Pressley, Elliot-Faust, and Miller (1985, p. 4) have defined a strategy as being,

“… composed of cognitive operations over and above the processes that are natural consequences of carrying out the task, ranging from one such operation to a sequence of interdependent operations. Strategies achieve cognitive purposes (e.g., comprehending, memorizing) and are potentially conscious and controllable activities.”

Strategies, therefore, are goal-directed behaviors that can be used to engage cognitive processes necessary for successful multiple document learning.

Question generation, or self-questioning, is one strategy that has the potential to support multiple document learning. Question generation has been effective in supporting single-text comprehension (Graesser & Lehman, 2011; Graesser, Lu, Olde, Cooper-Pye, & Whitten, 2005). In a study by Cameron, Van Meter, and Cameron (2018), for example, experimental participants were instructed on how to generate their own questions aimed at a multiple document comprehension. Participants who were able to generate a question aimed at multiple document comprehension performed better on a measure of multiple document learning than participants who did not generate a multiple document question.

It is important, however, to note that not all questions are equal. That is, there are
different questions for different purposes and which access different levels of knowledge (i.e., factual, surface-level processing vs. higher-order processing). A literature review by Chin and Osborne (2008) shows that learners often fail to generate questions that go beyond a factual understanding of documents. Cameron, Van Meter, and Long (2017), however, demonstrated that with instruction, learners are capable of generating questions aimed at higher-order processing and integration of multiple texts that could be used to support multiple document learning.

**Purpose of the Current Research**

The current research tests the effects of question generation strategy instruction on multiple document learning (Cameron et al., 2017, 2018). Cameron et al. (2017) developed Multiple Document Question Instruction (MDQI), a brief strategy instruction intervention that taught learners how to generate questions. Specifically, participants were taught to generate Multiple Document Questions (MDQs), which are questions aimed at multiple document integration and learning. Participants in the experimental condition received MDQI after reading four historical texts and generated questions based on the documents they read. Participants in the control condition read the same four historical texts and also generated questions but without specific questioning instruction. The MDQI condition generated significantly more MDQs than control condition peers. This study provided evidence that the intervention did support the learners' ability to generate MDQs.

In a follow-up study, Cameron et al. (2018) tested the effects of MDQI on multiple document learning. Participants were randomly assigned to either an experimental multiple document question instruction condition or a control condition (MDQI or NoQI), studied two text-based documents, and generated questions to help integrate those two documents after
reading. Participants in the experimental condition received MDQI after reading the texts. Following question generation, all participants were prompted to select a question they generated that, “BEST helps [them] to understand, evaluate, and synthesize the information from the multiple documents [they] just read.” All participants then wrote an essay response explaining the historical events and completed a multiple-choice posttest to assess factual knowledge gain. Results demonstrated that again, participants in the MDQI condition generated significantly more MDQs than control condition participants. There was not, however, a significant association between the condition and question selection. Finally, participants who generated an MDQ also performed better on the multiple document learning essay measure regardless of condition assignment. This result suggests that although college students are able to recognize MDQs, instructional support is needed for MDQ generation.

On the whole, the results of Cameron et al. (2018) demonstrate that the generation of MDQs can support multiple document learning, but the direct effects of question instructions were relatively small. A possible explanation for these findings is that experimental participants in this prior research did not receive MDQI until after reading the texts. The MD-TRACE model suggests that an internal representation of the documents knowledge is constructed during reading. Based on this theoretical model, question training may have a greater impact on multiple document learning if experimental participants are provided with MDQI before reading to allow them to begin generating questions while reading, as opposed to receiving question instruction and generating questions after reading. MDQI before reading the documents could potentially direct the learner's attention to multiple documents and their use to influence their approach to reading. The current study replicates that of Cameron et al. (2018) but includes a third condition in which participants receive MDQI before reading the documents.
This current study includes two experimental and one control condition. Participants in the first experimental condition, MDQI Before Reading (MDQI-BR), receive the question generation instructions before reading. In the second experimental condition, MDQI After Reading (MDQI-AR) participants receive the question generation instructions after reading. Control condition participants are told to study the documents carefully. All participants read the same set of six historical documents. After reading, all participants were directed to first generate questions and then to select their best question to understand, evaluate, and synthesize information from the multiple documents they read. Following question generation and selection, all participants wrote an essay response in which they explained the historical events that were described in the documents and then completed a multiple-choice posttest to assess factual knowledge gain. The essay responses were scored for both the number of perspectives included in the response and the textual evidence included in the response.

The first set of hypotheses in this study examine the effects of condition on question generation, question selection, and learning. Hypothesis one predicts that there will be a significant association between condition and the quality of questions that learners generate. That is, participants who receive MDQI instruction are expected to generate more MDQs and fewer lower-level questions than control condition participants. This pattern would replicate the findings of Cameron et al. (2017, 2018). In addition, it is expected that MDQI-BR participants will generate more MDQs than MDQI-AR participants.

Hypothesis two predicts that there will be a significant association between condition and the selection of a multiple document question. In previous research, it was predicted that participants in the question instruction condition would not only generate more MDQs but would also be better able to identify an MDQ than their peers in the control condition (Cameron et al.,
Contrary to this prediction, results indicated that participants were able to identify questions aimed at multiple document comprehension with or without the question instruction intervention. The current research will test this hypothesis again with the three conditions to see if there is a significant association between condition and question selection.

Hypotheses three and four examine the effects of condition on multiple document learning through the measures of multiple document integration and document knowledge. In previous research, experimental participants did not receive MDQI until after reading the texts. In that study there were no MDQI effects on learning. The current study replicates that of Cameron et al. (2018) but includes a third condition in which participants receive MDQI before reading the documents. Hypothesis three predicts there will be a significant association between condition and multiple document integration. That is, participants in the MDQI-BR condition are expected to have higher scores of multiple document integration than the MDQI-AR and control condition. Following this logic, hypothesis four predicts that there will be a significant difference between conditions on document knowledge. MDQI-BR participants are expected to score higher on a measure of document knowledge than MDQI-AR participants and control condition participants.

The second set of hypotheses examine the effects of self-generated MDQs on learning. Previous research found that those who generated an MDQ performed better on measures of multiple document learning than those who did not generate an MDQ (Cameron et al., 2018). This research aims to replicate these effects. Hypothesis five predicts a significant association between those who generate an MDQ and multiple document integration. Participants who generate an MDQ regardless of condition are expected to have higher scores of multiple document integration than those who do not. Hypothesis six predicts that there will be a
significant difference between those who generate an MDQ and those who do not generate an MDQ on document knowledge. Participants who generate an MDQ are expected to have higher posttest scores on average than participants who do not generate an MDQ.
Chapter 2

Literature Review

Introduction

Learning from multiple documents has become a critical skill for the modern learner (Goldman & Scardamalia, 2013). Despite the importance of learning from multiple documents (Nokes, Dole, & Hacker, 2007), research consistently demonstrates that learners struggle with these tasks (Britt & Sommer, 2004; Wineburg, 1991). Fortunately, instructional strategies can support and improve multiple document learning (e.g., Anmarkrud, Bråten, & Strømsø, 2014). The current study tests the effects of Multiple Document Question Instruction (MDQI; Cameron et al., 2017), a brief strategy instruction intervention designed to teach learners how to generate high-quality multiple document questions (MDQs). This instruction is intended to increase the number of MDQs learners generate, improve the ability to recognize an MDQ, and increase attention to multiple texts to improve multiple document learning. Cameron et al. (2018) tested the effects of this instruction, but experimental participants in that research did not receive MDQI until after reading the texts. Although some positive effects of instruction were found, these effects were small. The current study replicates that of Cameron et al. (2018), but includes a third condition in which participants receive MDQI before reading the documents. This chapter outlines the theoretical and empirical research that underpins this study.

Theoretical Framework

MD-TRACE. The theoretical model that explains multiple document learning is the Multiple Document Task Relevance and Content Extraction model (MD-TRACE, Rouet & Britt, 2011). MD-TRACE intends to describe the ideal process of multiple document reading comprehension and learning. It accounts for both internal and external factors to the learner.
during a multiple documents learning task and the interaction of these factors throughout the process.

MD-TRACE explains that the presentation of a task first initiates the multiple document reading process by providing a purpose for reading. The learner creates a task model, an internal representation of the task and its criteria. In order to be successful, the task model must include multiple documents. The task model guides the learner towards what to read, what information to select, and how to relate the documents to one another. The learner reads through the documents and selects what he or she believes to be relevant information for the completion of the task. The learner will continue to work through these documents in this recursive manner, selecting and organizing information from the documents, until he or she achieves the task model objective. The learner will then compare the task product, the physical manifestation of the task model, to the learner's standards of the task model. The result is both an external task product and an internal structure of knowledge that represents the multiple documents learning that took place.

**Theoretical principles and assumptions.**

*The Construction-Integration Model.* Kintsch’s (1988) Construction-Integration Model explains the underlying processes of comprehending documents and is the theoretical underpinning for the MD-TRACE model. Kintsch explains that there are three levels of text comprehension that are internal representations of comprehension for the learner: the textbase, the proposition representation, and the situation model. The textbase is a verbatim copy of the printed text. The proposition representation is the understanding of the relationship among words in the written sentence, providing meaning to the textbase. Finally, the situation model provides context and meaning to what is being read and comprehended at the textbase and propositional
levels. Comprehension takes place at all three levels through two interacting processes, construction and integration.

*The Documents Model.* The Documents Model (Perfetti et al., 1999) takes Kintsch’s Construction-Integration Model one step further by explaining how multiple documents are comprehended and integrated. The Documents Model proposes that as the learner reads, the learner selects information to build a mental representation of the fundamental concepts from each document and relates these concepts to other documents read. This document information is also organized and stored along with the sourcing information. The learner begins to build connections across documents by recognizing how the documents’ content and sourcing information relate to each other. These connections across documents are part of the Situations Model, which provides context across multiple documents. The documents and situations models serve as the theoretical mental representations of multiple document comprehension and integration from which we can assess what learning has occurred.

*Elements of reader processing.* The multiple document reading process is also affected by factors both internal and external to the reading process (Brit & Rouet, 2011). Internal factors include prior knowledge, prior beliefs about a topic, epistemological beliefs, reading ability, strategy use, and self-regulation. These internal factors account for the individual differences that a learner brings to a multiple documents learning situation. These factors interact with the learner’s ability to select information from the text, organize information within and among texts, and make connections to prior knowledge. External factors include the documents and the documents’ characteristics, sources, the task, and the domain of the document materials.

*Internal factors.* Many internal factors contribute to multiple documents learning. These internal factors are individual difference variables that contribute to multiple document
comprehension, processing, and integration. These factors include prior knowledge, epistemological beliefs, prior beliefs about conflicting topics, reading abilities, strategies, and self-regulation (Rouet & Britt, 2011). These internal factors interact with the external factors that contribute to the overall ability of a learner to read, process, select information, and assess and adjust their final task product to meet the task demands.

External factors. There are also many external factors such as the task, the documents, the content of the documents, and the documents’ characteristics, that can impact multiple document learning. For example, the document set may have documents with conflicting content that the reader must try to understand and reconcile that would affect the way in which the learner processes the documents. Additionally, the documents could interact with the internal factors such as individual beliefs to influence what content the learner attends to, selects, and includes in their final mental representation.

Research on Learning from Multiple Documents

Although the modern learner often engages with multiple documents in both academic and non-academic settings, researchers have found that students often struggle to develop an integrated understanding of the information presented across sources (Wiley & Voss, 1999; Britt & Sommer, 2004). As described in the MD-TRACE model, learning from multiple documents is a cognitively demanding task that requires a complex interplay between internal and external resources and the learner’s ability to regulate themselves throughout the entire process (Rouet & Britt, 2011; Bråten et al., 2013). Given this, it is not surprising then that learners struggle with these tasks. This level of multiple document learning requires the coordination of strategic processing to engage in the cognitive processes of selecting relevant document information,
organizing document information, and integrating document information (Bråten, Anmarkrud, Brandmo, & Strømsø, 2014).

**Defining a document.** A document can be defined as a source of information on a particular topic. This definition is broad because documents can vary in their content in several different ways. First, documents can represent information from different disciplines such as science or history. With these different disciplines comes different conventions for presenting information. For example, a document in science may be an experiment that was conducted to demonstrate the effects of climate change. For history, a document may be a diary entry accounting for a day in the life of a civil war soldier. Documents can even vary within a discipline, and many studies often include information on the types of documents used in the research such as magazine articles, online news articles, or textbooks (e.g., Stadtler & Bromme, 2007). Documents can also contain or be comprised of different types of representations (Van Meter & Cameron, 2018). Documents can consist of only text or can be just a visual representation such as a photograph. Documents can also contain both text and visual information within a single document.

The purpose of research in multiple documents learning is to examine how learners interact with document sets. Documents are an external factor of the MD-TRACE model (Britt & Rouet, 2011). A document set can contain a variety of document types described above (e.g., Strømsø, Bråten, & Britt, 2010). A document set can represent different subject areas such as science, evaluating a scientific phenomenon like climate change (e.g., Hagen, Braasch, & Bråten, 2014) or making health decisions like whether or not to vaccinate your child (e.g., Maier & Richter, 2013), or history, evaluating what actually occurred during a historical event (e.g., Britt & Sommer, 2004; Wiley & Voss, 1999). What is particularly important to the learning process,
as represented in the MD-TRACE model, is the nature of the relationships of the documents within a document set. Document sets can contain both complementary or conflicting information (e.g., Bråten & Strømsø, 2010; Bråten, et al., 2014; Nokes et al., 2007). Within the document set on climate change, for example, there may be documents that support or contradict the claim that the planet is experiencing a marked change in global temperatures. How the learner uses this complementary or conflicting information is relevant to the task demands and requires strategic processing (Bråten et al., 2014). The current study includes both text and visual documents to approximate the conditions of an authentic multiple documents learning task in history.

Measurement of Multiple Documents Learning

The measurement of multiple document learning is not straightforward due to the complexity of both the task and the knowledge that may be acquired. MD-TRACE demonstrates that learning happens at a singular document level and a multiple document level. First, singular document learning can occur at a surface-level, memory of the text or could require the depth of processing that connects the document to a learner’s prior knowledge. Second, this learning does not just take place for one document, but for multiple documents. This requires the understanding of the relationships between documents, the evaluation of sources, and integration of knowledge. Multiple document learning can therefore be evaluated on each of these levels by examining the learner’s change in knowledge or their learning process. Not surprisingly, there are a variety of measurement techniques in this body of work. Specifically, researchers have focused on approaches that either examine the process (e.g., Bråten & Strømsø, 2003; Cerdán & Vidal-Abaraca, 2008) or the product (e.g., Bråten, Strømsø, & Britt, 2009) of multiple document learning. The three most commonly used measures in multiple document research have been
think alouds (a process measure), and verification tasks and essays (product measures). While the measurement of process and product are different, they each provide insight into what is learned and how it is learned. Together, these data inform instruction and measurement of multiple documents learning.

**Process measures of multiple documents learning.** Process measures capture the thoughts of the learner while reading and processing multiple documents. These measures show decision points, attention, and self-regulatory processes (Bråten et al., 2014). Process measures can also reveal what strategies learners use, when they use them, and why they use them (Bråten & Strømsø, 2003).

**Think-aloud protocols.** Think aloud protocols measure the online processing of textual information (Pressley & Afflerbach, 1995). As described previously, this format gives insight into the online processing of what learners attend to, how they interpret the task, and how they regulate themselves throughout the task. This information provides a more complete understanding of the multiple document learning process beyond the final task product (Anmarkrud et al., 2013; Bråten et al., 2014; Bråten & Strømsø, 2003). However, think-aloud protocols are demanding of time and space, and for the learner, require considerable cognitive effort. As Cerdán and Vidal-Abaraca (2008) demonstrate, think-aloud protocols can hinder multiple text processing. Verbal protocols take up significant working memory space for verbal processing that could be dedicated to multiple documents learning.

**Product measures of multiple document learning.** The product is a glimpse into the internal knowledge structure that the learner has created, the Documents Model (Perfetti et al., 1999). The product gives the researcher a momentary snapshot of what the learner has attended
Verification tasks. Verification tasks have been used to measure the product of multiple document learning. There are three types of verification tasks intended to measure each level of text representation (Kinstch, 1988). Sentence Verification Tasks (SVT), developed by Royer, Carlo, Dufresne, & Mestre (1996), measure the textbase knowledge of a learner for a single document. The measure consists of four types of items from sentences in the experimental text (1) original sentences, (2) paraphrases, (3) meaning changes, and (4) distractors (Bråten, Strømsø, & Samuelstuen, 2008). Participants indicate items with the same meaning as the original text with a “yes” and items with a different meaning as the original text with a “no.” The Intratext Inference Verification Task (IntraIVT) measures the relationships among information in the text, also known as the situation model of a single text. For this measure, the participants indicate if an item contains valid or invalid inferences from the text. Participants indicate if the item inference was valid or invalid. Finally, the Intertext Inference Verification Task (InterIVT) measures the relationships among information across multiple documents. For this measure, participants indicate if the item contains valid or invalid inferences from multiple texts. The InterIVT measures the Documents Model. Together, a research team can infer what knowledge a learner has at the textual, situation, and documents model levels of reading comprehension. Also, Stadler et al. (2013) developed a measure called the Conflict Verification Task (CVT) to measure the conflicting intertextual relationships among the documents. The CVT measures the intertext model representation and influence of sources on conflict memory.

These tasks measure knowledge at each level of representation, which directly aligns with the theoretical cognitive representation of a product of multiple documents reading. This
theoretical support lends itself to the validity argument of this measure. While Bråten and Strømsø (2010) have explained that the reliability estimates have reached levels of acceptance by the measurement community for researcher developed tools, many of the studies have reported weak reliability estimates in the .50-.60 range. These low-reliability estimates have been consistent across several studies (i.e., Bråten & Strømsø, 2006, 2010, 2011; Bråten, Strømsø, & Britt, 2009; Strømsø, Bråten & Britt, 2010). New items for these measures must also be developed based on the texts used in each study, which may also be contributing to this issue of low reliability.

**Essays.** Essays provide a product of the learners’ internal knowledge structure based not on recognition like the verification tasks, but from recall and generation. Most commonly in multiple document research, researchers have selected structured essay prompts over free-recall prompts to measure content knowledge, the organization of that knowledge, and the connections between content knowledge and source information. (Anmarkrud et al., 2014; Kobayashi, 2009, 2010, 2014; Le Bigot & Rouet, 2007; Stadtler et al., 2013).

The flexibility of the essay also lends flexibility to the scoring of these essays. For example, Gil and colleagues (2010) measured the transformation of textual information to measure deep processing versus superficial processing (Perfetti et al., 1999). Others have measured the strength of an argument using a rubric developed by Reznitskaya, Kuo, Glina, & Anderson (2009) with better arguments being constructed by considering both sides of an argument and providing claims, reasons, and evidence. This variety of ways to think about the way in which we can conceptualize what an integrated multiple document essay looks like could limit the conclusions that can be drawn across the broad body of research.
**Conclusions.** Think alouds, verification tasks, and essays all offer unique insight into the processes and products of multiple documents learning as described above. Each has its strengths and its weaknesses. However, our definition of multiple documents learning means something slightly different for each of these measurement approaches. For essays, it could be the integration of content to content or document to document (e.g., Le Bigot & Rouet, 2007), the creation of an argument (e.g., Kobayashi, 2009), or the transformation of knowledge (e.g., Gil et al., 2010a). For verification tasks, it is the depth of processing and text memory (e.g., Bråten & Strømsø, 2011; Bråten, Strømsø, & Samuelstuen, 2008). For think alouds, it is strategy use and self-regulation (e.g., Bråten & Strømsø, 2003) or processing of text in the selection of content (e.g., Anmarkrud et al., 2013).

In the current study, multiple document learning was measured by an essay prompt that asked participants to explain a historical conflict using evidence from six conflicting documents. These essay responses were scored based on the holistic rubric that measures the integration of multiple perspectives and the accuracy of textual evidence. This rubric, which was developed by Rukavina and Daneman (1996), has been in multiple document used in previous studies (e.g., Bråten et al., 2013). Bråten et al. (2013) and Rukavina and Daneman (1996) designed their essay prompts to consider the claims and reasons of two conflicting perspectives presented within the document sets. The rubric, therefore, was designed to measure how well participants could comprehend and evaluate these claims and reasons through their response, and was used as the dependent measure in these studies. First, the researchers considered whether or not the content of the responses was accurate based on the information presented within the document set. Next, they evaluated the responses for the number of perspectives presented in the response. Then researcher determined whether or not those perspectives were supported with textual evidence.
from the document set. The essay prompt in this study follows the same standards in prompting the examination of the perspective claims and reasons and this rubric has been adapted to evaluate the content of this study and described further in Chapter 3.

Participants were also evaluated on document knowledge using a multiple-choice measure developed by the researcher. This measure was intended to measure document information recall. The intent in using this measure over a verification task is to target what documents learners attend to and what information they comprehend. Consistent with the previous research reviewed above, participants were not able to return to the documents during the writing or multiple-choice tasks (e.g., Anmarkrud et al., 2014; Bråten et al., 2013; Gil et al., 2010a; Le Bigot & Rouet, 2007; Strømsø et al., 2010). Participants were allowed to take notes while they study the documents and these artifacts of multiple documents processing assisted in interpreting the essay responses and document knowledge posttest results, as in previous research (e.g., Hagen et al., 2014; Kobayashi, 2014).

**Supporting Multiple Document Learning**

Multiple documents research has focused on understanding how the manipulation of tasks can impact student learning from multiple documents. The search for the ideal task to promote multiple document learning, however, has yielded mixed results. Some studies found that argument tasks outperformed summary tasks on multiple documents learning (Britt & Sommer, 2004; Gil et al., 2010a; Gil et al., 2010b; Hagen, Braasch, & Bråten, 2014; Wiley & Voss, 1999) while others found that summary tasks outperformed argument tasks on multiple documents learning (Gil et al., 2010a; Gil et al., 2010b; Le Bigot & Rouet, 2007). These mixed results are likely a product of interactions between the definition of the task, the documents, and
individual differences such as prior knowledge, epistemological beliefs, and strategy use (Bråten et al., 2014).

While the task is something that the learner cannot always control as an external factor, the learner can control his or her reading behaviors and the proper use of strategies to support learning throughout different tasks. This is why it is essential to study strategies that can generalize across different multiple document tasks and develop instruction that supports both the how and when to use these strategies.

**Strategy use and multiple document learning.** In a seminal study, Wineburg’s (1991) sought to understand the differences between experts and novices learning from multiple historical documents. He found these differences between experts and students were the heuristics experts use. These heuristics, which Wineburg (2001) later labeled as “Historical Thinking,” include sourcing, contextualization, and corroboration. Sourcing is the act of considering authorship, time period, and biases of the materials. Contextualization is the act of placing an event in space and time to understand the content described in the document. Lastly, corroboration is using multiple documents to derive intertextual meaning by comparing and contrasting those materials. Although Wineburg's primary purpose was understanding multiple document learning in the discipline of history, these heuristics are applicable across disciplines that use multiple document learning. These heuristics would serve a broader purpose if they were labeled “Multiple Document Thinking,” and this will be referred to as such throughout the remainder of this document. These heuristics are examples of the strategic processing that takes place during successful multiple document learning.

The MD-TRACE model supports the importance of strategy use in learning from multiple documents (Rouet & Britt, 2011). Strategies are considered to be an internal factor that
learners bring to a multiple documents task. The task helps the learner to determine what strategies to employ during the learning process. Researchers have studied the effects of strategy use through the examination of online processing of multiple documents. Anmarkrud, McCrudden, Bråten, & Strømsø (2013), for example, conducted a study in which they instructed participants to give advice to a friend concerned about cell phone use and health risks. Participants studied six documents while thinking aloud and then responded to the essay prompt to give their friend advice based on what they read. Think aloud data was coded for the presence of strategy use such as the verbal judgments of more- or less-text relevant information and source evaluation. They found a positive correlation between essay quality (their measure for multiple document integration) and verbal judgments at more text-relevant and less text-relevant sections of the texts. They also found a significant correlation between essay scores and participants’ use of source evaluation. The authors interpreted these utterances as an indication that participants were strategically sorting through the information they needed and the information they did not need to construct their understanding of the text.

In another study that used think alouds, Bråten and Strømsø (2003) studied law students as they worked through their course materials and measured their elaborations, memorization, problem-solving, and comprehension confirmation. These think-aloud sessions took place over three separate occasions as participants learned course materials in preparation for a class exam. They found that the depth of processing, time on task, and strategy use varied with the task specifications as learners went from reading the course materials for the first time to reading the materials to review for the exam. The selection of a suitable strategy is dependent on the learner’s procedural and conditional strategy knowledge. For example, when one participant read during his first pass, he read in a linear manner and took few notes, but in preparation for the
exam, he took notes and frequently worked back-and-forth between the textbook and code of law book using elaboration and self-monitoring strategies. In contrast, another participant used the same approach from the first pass to the preparation for the exam, reading in a linear fashion with little note taking. These examples illustrate how the first participant described knew how and when to employ reading strategies to support preparation for the exam and could adjust according to the demands of the task. The second participants, however, did not demonstrate that same procedural and conditional strategy knowledge despite changes in the task. In addition, there was evidence that learners shift their reading behaviors based on individual differences such as prior knowledge. Participants with related domain knowledge were able to construct more textual links that supported learning during reading. Such links led to stronger performances on the exam.

Each of these studies lends further support to the need for strategic processing for successful multiple document learning. They also illustrate how the interactions between both internal and external factors impact strategic processing behaviors. It is important to note, however, that each of the previous studies only provide insight into what students do without intervention. It is possible that strategy instruction could improve strategy selection and use in multiple document learning, but few studies have examined this opportunity. The primary purpose of this research is to examine the effectiveness of question generation as a multiple documents reading comprehension strategy.

**Question Generation and Multiple Document Learning**

Question generation has been demonstrated to be an effective reading comprehension strategy in single-document comprehension (Graesser, Lu, Olde, Cooper-Pye, & Whitten, 2005; Rosenshine, Meister, & Chapman 1996). The strategy provides the learner with the opportunity
for comprehension monitoring and better self-regulation than goal setting (Graeser & Lehman, 2011). Additionally, studies have found that questions can help direct attention to specific information that the learner must select to answer the question (Cerdán & Vidal-Abarca, 2010; Taboada & Guthrie, 2006). Still, the potential benefit is dependent on the quality of the question generated (Chin & Osborne, 2008).

Question strategies have proven to be effective in multiple-document comprehension as well. For example, in a study by Britt and Sommer (2004) undergraduate students studied multiple conflicting history texts for either comprehension or integration. The participants completed an intervening task, either micro-structure questions (focusing on small micro details of the text) or macro-structure questions (focusing on the why and what happened). Participants in the macro-structure question condition performed better than micro-structure question peers on integration, measured by the number of switches between source information in the written response. Similarly, in a study by Taboada and Guthrie (2006), elementary students studied ecology texts and generated questions based on the texts they read. Participants who generated higher-level questions, questions aimed at cross-text integration, demonstrated better conceptual understanding and reading comprehension.

Both studies demonstrate that for a question to support multiple document learning, the question must be a higher-order question aimed at inter-textual connection building. That is not to say that factual, lower-order questions do not serve a purpose in multiple document learning. Instead, the learner must go beyond this lower-order question and seek to build inter-textual connections. However, college-students often struggle to generate high-quality questions without instructional support as consistently demonstrated in the research (Cameron et al., 2017; Cameron et al., 2018; Chin & Osborne, 2008). For this reason, Cameron and colleagues (2017)
developed and tested the Multiple Document Question Instruction (MDQI) intervention to support Multiple Document Question (MDQ) generation with positive results. MDQI was designed to teach learners what MDQs are and how question generation can be used to help the learner to understand and learn from multiple documents. Instruction included two parts: a video to explain what MDQs are and how to use them to support learning and MDQ frameworks that provided examples of MDQs. The current study tests the effects of MDQI on multiple document question generation. The study also tests order effects of MDQI because questions can help to direct attention to relevant document information for selection (Cerdán & Vidal-Abarca, 2010; Taboada & Guthrie, 2006). Experimental participants received MDQI before reading or after reading the documents and were compared on the quality of questions generated.

The Current Study

The current study is the next step in the assessment and development of the MDQI. Previous research (Cameron et al., 2018) demonstrated that there was a condition effect on question generation, but not multiple document learning. There was, however, an effect on learning for those who generated MDQs. Based on these results, question training may have a more significant impact on multiple document learning if participants receive MDQI before reading to allow them to generate questions while reading. In this current study, participants were assigned to one of three conditions: MDQI Before Reading (MDQI-BR), MDQI After Reading (MDQI-AR), or control. Participants received instructions relative to their assigned condition to study six documents. After studying the documents, participants generated questions based on those documents and selected their best question. Following question generation and selection, participants were asked to respond to an essay prompt and a multiple-choice document
knowledge posttest. Researchers scored these essay responses for the inclusion of multiple perspectives and for textual evidence in those responses.

Additionally, participants completed individual difference measures including reading comprehension and prior knowledge at the beginning of the study, before the experimental and control materials. It is expected that there are no significant differences between the three conditions in reading comprehension abilities and prior knowledge. Two one-way ANOVAs will test this assumption.

The first set of hypotheses in this study examine the effects of condition on question generation, question selection, and learning. Hypothesis one predicts that there will be a significant association between condition and the quality of questions that learners generate. A 3 x 5 (condition x question quality category) chi-square test will evaluate if participants who receive MDQI instruction are expected to generate more MDQs and fewer lower-level questions than control condition participants. This pattern would replicate the findings of Cameron et al. (2017, 2018).

Hypothesis two predicts that there will be a significant association between condition and the selection of a multiple document question. In previous research, it was predicted that participants in the question instruction condition would not only generate more MDQs but would also better able to identify an MDQ than peers in the control condition (Cameron et al., 2018). Contrary to this prediction, results indicated that participants were able to identify questions aimed at multiple document comprehension regardless of instruction condition. The current research will use a 3 x 2 (condition x MDQ selection) chi-square analysis to test this hypothesis to see if there is a significant association between condition and question selection.
Hypotheses three and four examine the effects of condition on multiple document learning through the measures of multiple document integration and document knowledge. In previous research, experimental participants in that research did not receive MDQI until after reading the texts. Although some positive effects of instruction were found, these effects were small. The current study replicates that of Cameron et al. (2018) but includes a third condition in which participants receive MDQI before reading the documents. Question training may have a greater impact on multiple document learning if experimental participants are provided with MDQI before reading to direct attention to multiple document information while reading, rather than reflecting on document information after reading. Hypothesis three predicts there will be a significant association between condition and multiple document integration. A 3 x 6 (condition x essay integration category) chi-square analysis will test this hypothesis. Participants in the MDQI conditions are expected to have higher scores of multiple document integration than the control condition, and the MDQI before reading condition will score highest on multiple document integration. Following this logic, hypothesis four predicts that there will be a significant difference between conditions on document knowledge. MDQI participants are expected to score higher on a measure of document knowledge than control condition participants, with MDQI before reading participants scoring the highest on document knowledge. A one-way ANOVA will test the effects of condition on document knowledge posttest scores.

The second set of hypotheses examine the effects of self-generated MDQs on learning. Previous research found that those who generated an MDQ performed better on measures of multiple document learning than those who did not generate an MDQ (Britt & Sommer, 2004; Cameron et al., 2018; Cerdán & Vidal-Abarca; Taboada & Guthrie, 2006;). This research aims to replicate these effects. Hypothesis five predicts a significant association between those who
generate an MDQ and multiple document integration. A 2 x 6 (MDQ x essay integration category) chi-square analysis will test this hypothesis. Participants who generate MDQs are expected to have higher scores of multiple document integration than those who do not.

Hypothesis six predicts that there will be a significant difference between those who generate an MDQ and those who do not generate an MDQ on document knowledge. Participants who generate an MDQ are expected to have higher posttest scores on average than participants who do not generate an MDQ. A one-way ANOVA will test the effects of MDQ generation on document knowledge posttest scores.
Chapter 3

Method

Participants

Participants were 109 students recruited from an introductory undergraduate Educational Psychology course. Participants were offered course extra credit for the completion of the study. The participants in this study were predominately female (94%), Caucasian (95%), and spoke English as a first language (99%). Approximately 64% of participants were Education majors from a variety of concentrations; (e.g., Secondary Science Education, Elementary Education). The second most common major represented in this sample was Communication Sciences and Disorders, at 28%. The remaining 8% of the sample included participants who were undecided on a major or enrolled in Psychology or Science majors. The average reported GPA of the sample was 3.40. About 75% of the participants were in their first year of undergraduate education, and 21% were in their second year. The majority of participants (69%) had taken three or fewer credits or between 4-6 credits (14%) of college history, at mostly introductory levels (89%). These figures indicate that participants in this study had little advanced academic preparation in history. Participants consented to participate during the experimental session using an electronic informed consent form.

Design

The current study tests the effects of Multiple Document Question Instruction (MDQI) on multiple document learning in history with three separate conditions. The Qualtrics survey system (Qualtrics; Provo, UT) system randomly assigned participants to the MDQI Before Reading (MDQI-BR), MDQI After Reading (MDQI-AR), or Control condition. All participants completed the informed consent, demographic survey, reading comprehension measure (GRE;
Participants received instructions relative to their condition and studied a historical topic document set. Participants in the MDQI-BR condition received the instructional components of the MDQI before reading and studying the document set. Participants in the MDQI-AR condition received the instructional components of the MDQI after reading and studying the document set. The Control participants were instructed only to study the materials carefully before going on to complete the dependent measures of the study. All participants generated questions based on the documents after studying the documents. After question generation, participants were also directed to select their best question. Following the question generation and selection process, all participants generated a response to the essay prompt and answered a 10-item knowledge posttest to measure multiple document learning.

Materials

**Individual Difference Measures.**

*Demographic survey.* This survey asked participants to report their gender, ethnicity, SAT, ACT, current GPA, academic year (i.e., Freshman, Sophomore, etc.), first language, and intended major. The demographic survey is located in Appendix A.

*Reading comprehension.* The reading comprehension measure is a 9-item reading comprehension measure from the Graduate Record Exam (GRE; Reading Comprehension Sample Questions, 2014). This measure consisted of 6 items that assessed comprehension of short reading passages and three vocabulary cloze task items. These vocabulary cloze task items were included because the understanding of key vocabulary terms also requires the comprehension of the text in which those words are being used and has served as a valid measure of reading comprehension in recent studies (e.g., Alexander et al., 2016). Items are in Appendix
A. Participants in the Alexander et al. (2016) study were primarily sophomores and juniors at a large mid-Atlantic university enrolled in human development courses. The majority of these participants (59%) were behavioral science majors. The reliability estimate of the scores from the Alexander et al. (2016) study was acceptable (Cronbach’s alpha = .79). Despite implementing this measure in the same manner and with a similar population as the previous research, the reliability estimate of these scores was lower (Cronbach’s alpha = .28). However, the mean and standard deviation ($M = 2.92$, $SD = 1.55$) were similar to that of the previous sample ($M = 3.21$, $SD = 1.62$) (Alexander et al., 2016). Scores were used to compare the reading comprehension abilities of condition groups. The explanation and implication for the poor reliability of this measure is discussed further in the results and discussion sections.

**Prior knowledge.** A 10-item multiple-choice test measured the participants' prior knowledge of American Imperialism. This test included sample items from the NY Regents exam and has been used in prior research (Cameron et al., 2017; 2018). Despite implementing this measure in the same manner and with a similar population as the previous research, the reliability of scores in the current study was low (Cronbach’s alpha = .25). The mean was also one point lower than previous research ($M = 4.40$, $SD = 1.71$) indicating a lower level of prior knowledge and a higher possibility for guessing on items, which could have impacted the reliability of the scores. Scores were used to compare prior knowledge of condition groups. The explanation and implication for the poor reliability of this measure is discussed further in the results and discussion sections.

**Document Knowledge Posttest Measure.** A 10-item multiple-choice posttest assessed participants’ knowledge of the relevant document information from each document. This measure consisted of items that assessed knowledge of the visual and text-based experimental
documents. The measure determined what information participants selected from the individual documents and from which types of documents learners were selecting and comprehending information. The measure also assessed the knowledge that would be acquired from bridging information from across multiple documents in the document set. Table 1 contains a blueprint of the measure with regards to what types of documents, how many, and which documents were necessary to answer the item. The items are in Appendix A. The reliability estimate of these scores for the full measure was low (Cronbach’s alpha = .50). Additionally, the subtest score reliability estimates were also low for both visual document subtest scores (Cronbach’s alpha = .43) and text-based document subtest scores (Cronbach’s alpha = .43). Analyses of these posttest scores are further described in the results and discussion sections.

Table 1

*Document Knowledge Posttest Measure Item Composition*

<table>
<thead>
<tr>
<th>Item</th>
<th>Type</th>
<th>Single or Multiple</th>
<th>Text 1</th>
<th>Text 2</th>
<th>Photo 1</th>
<th>Photo 2</th>
<th>Political Cartoon 1</th>
<th>Political Cartoon 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Text</td>
<td>Single</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Visual</td>
<td>Single</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Text</td>
<td>Single</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Text</td>
<td>Single</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>5</td>
<td>Text/Visual</td>
<td>Multiple</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>6</td>
<td>Text</td>
<td>Multiple</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Visual</td>
<td>Single</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>8</td>
<td>Text</td>
<td>Multiple</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Visual</td>
<td>Multiple</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>10</td>
<td>Text</td>
<td>Multiple</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Instructional Materials.

Documents. Participants received a set of six historical primary source documents electronically via the Qualtrics survey system. Documents were selected from both an undergraduate history textbook commonly used at the university, entitled *Major Problems in the Gilded Age and Progressive Era: Documents and Essays* (Fink, 2001), and a historical archive website used for instructional materials. Descriptions and sources of the documents are in Table 2 and copies of the documents can be found in Appendix B. As a set, the six documents presented conflicting information and opposing views on American Imperialism of the Philippines. There were three documents (i.e., text, photograph, and political cartoon) opposing the invasion of the Philippines and three documents (i.e., text, photograph, and political cartoon) defending the invasion of the Philippines. Each document offers unique content to trace the information contained in the participants’ responses back to its original source.
Table 2

*Description of Experimental Documents*

<table>
<thead>
<tr>
<th>Document</th>
<th>Title</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text 1</td>
<td>William Jennings</td>
<td>• Argues that the U.S. should not be involved with occupying the Philippines</td>
</tr>
<tr>
<td></td>
<td>Bryan Opposes U.S. Occupation of the</td>
<td>• 289 Words/ Flesch-Kincaid Grade Level 12.8</td>
</tr>
<tr>
<td></td>
<td>Philippines, 1890</td>
<td></td>
</tr>
<tr>
<td>Text 2</td>
<td>Albert Beveridge</td>
<td>• Argues that the U.S. should occupy the Philippines</td>
</tr>
<tr>
<td></td>
<td>Defends U.S. Imperialism, 1900</td>
<td>• 817 Words/ Flesch-Kincaid Grade Level 8.8</td>
</tr>
<tr>
<td>Photograph</td>
<td>No Title</td>
<td>• Depicts American soldiers marching through the streets of the Philippines</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>with a giant American flag.</td>
</tr>
<tr>
<td>Photograph</td>
<td>Filipino Nationalists Work Toward</td>
<td>• Depicts Filipino Nationalist in a political office talking and completing</td>
</tr>
<tr>
<td>2</td>
<td>Independence</td>
<td>paperwork to work toward independence.</td>
</tr>
<tr>
<td>Political</td>
<td>No Title</td>
<td>• Depicts a pirate-like character clothed in the American flag holding a</td>
</tr>
<tr>
<td>Cartoon 1</td>
<td></td>
<td>sword with the words &quot;Imperialism&quot; across the blade. Under the feet of the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pirate are some smaller characters wearing names of the countries under U.S.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Imperialism.</td>
</tr>
<tr>
<td>Political</td>
<td>Showing Light to</td>
<td>• Depicts an American escorting a Filipino toward a bright lamp that sheds</td>
</tr>
<tr>
<td>Cartoon 2</td>
<td>the Filipino People</td>
<td>light into darkness on top of a pile of books with the words freedom,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>education, justice, and commerce.</td>
</tr>
</tbody>
</table>
**Documents matrix.** Participants received a documents matrix (Cameron et al., 2017) to record and organize any notes they had from the documents. This matrix contains the elements of multiple document thinking heuristics of sourcing, contextualization, and corroboration (Wineburg, 1991), as seen in Figure 1. The matrix contains categories for both sourcing and contextualization to help learners select and organize information from the documents. The Sourcing categories prompt the learner to consider the title, type of document, who created the document, and the point of view of the author. Contextualization categories prompt the learner to consider why the document was created, the main ideas of the document, and importance of the document. Corroboration occurs by comparing the document information across rows.

Researchers did not provide instruction on how to use the matrix, nor were the participants required to use it. Participants were told that, because they would not be able to return to the documents later in the study, that they could use the matrix to record and organize their notes as they saw fit as in previous studies (e.g., Anmarkrud et al., 2014; Bråten et al., 2013; Gil et al., 2010a). These matrices were analyzed as artifacts of learner processing (Hagen et al., 2014; Kobayashi, 2014) to provide insight to patterns in learning outcomes found in the results and discussion sections.
Experimental MDQI materials. Participants in both the MDQI Before Reading and After Reading conditions received the experimental instruction electronically in the Qualtrics environment. Cameron and colleagues developed and used this instruction in previous research (Cameron et al., 2017; Van Meter & Cameron, 2018). This instruction included a 5-minute instructional video about multiple document question generation and provided question frameworks, described further below. The purpose of this instruction was to facilitate the generation of Multiple Document Questions (MDQs) and how to use these types of questions to improve multiple document learning. All MDQI materials are in Appendix C.

Instructional video. A five-minute video explained what multiple document thinking is and the role this thinking plays in learning from multiple documents. The video consists of

---

**Figure 1.** Documents Matrix. A notetaking and organizational tool for participants to use while reading the documents.
PowerPoint slides with voice-over narration. The narrator first explains the importance of using multiple sources of information to get a complete understanding of an event and how questions can help to understand information across those sources. This is explained using an anecdote about professor using questioning trying to corroborate two stories from students who missed an exam using a question. The narrator goes on to explain how questions can facilitate multiple document thinking and the integration to understand multiple perspectives of an event. Next, the slides provide an example of an MDQ and a low-level, single-text question while the narrator explains the differences between the question types. This explanation encourages the learner to consider what information is required to answer each type of question. Further, the narrator asks the learner to reflect on the connections that are made between information when asking an MDQ versus a low-level, single-text question. Finally, the video summarizes the instructions by reiterating the importance of asking quality multiple document questions while studying multiple documents.

*Question frameworks.* Participants in the MDQI conditions were also instructed to use MDQ question frameworks (Rosenshine et al., 1996) in combination with their documents matrix to generate questions. These frameworks are not only examples of MDQs, but also provide a structure for comparing multiple document information (i.e., How do the points of view differ in document x and document y?). Participants could use these frameworks to help them generate questions to synthesize information across documents. (Provided before directions to generate questions.)

*Control materials.* In the control condition, participants received instructions to study the documents carefully.
**Question generation measure.** After reading, all participants were prompted to “generate questions that help to understand, evaluate, and synthesize the information from the multiple documents you read.” Participants could generate up to five questions. From previous research, participants generated about five questions on average and this also allowed for a restricted range for question selection options (Cameron et al., 2017). Below the directions were five separate text boxes provided for students to type in their question generation. The author and another member of the research team coded each question using the question rubric described below. This rubric categorized the types of questions generated.

**Question selection measure.** For question selection, the questions that each participant generated populated on the following screen in a multiple choice-like format. The participant could then select a single question from the multiple choice format using the criteria of the prompt. The prompt read, “Please select the question that BEST helps you to understand, evaluate, and synthesize the information from the multiple documents you just read.” The key words “understand, evaluate, and synthesize” are the same key words used in the question generation prompt that was used for participants in all three conditions. This prompt was also consistent with the language used in the MDQI instruction that described an MDQ. For the analysis, all questions were coded using the question coding rubric described below. The selected question was dichotomously scored, (1) if the question selected was an MDQ and (0) if the question selected was not an MDQ.

**Document integration essay measure.** Participants responded to the prompt, “Write a letter to a friend explaining the conflict over the occupation of the Philippines and American imperialism in the late 19th and early 20th century. Support your response with evidence from the documents.” Participants entered their response in a text box below this prompt in the
electronic survey system. The author and another member of the research team coded each question using the document integration rubric described below.

**Procedure**

The author recruited participants during the scheduled class time. The author explained the study and that students would receive course extra credit for completion of the study. An alternative assignment was also available and explained. Participants used Sign-up Genius to register for an experimental session held in an on-campus computer lab. During the session, participants received a website code that was used to access the Qualtrics survey system via a designated website code that the author distributed to the participants. Participants then read and completed the informed consent, demographic survey, reading comprehension measure (Alexander et al., 2016), and prior knowledge measure. The Qualtrics survey system randomly assigned participants to one of three conditions: the MDQI Before Reading condition, the MDQI After Reading condition, and the control condition. Table 3 below depicts the order of study activities for each condition. Participants who were assigned to the MDQI Before Reading condition watched an instructional video, studied the documents, viewed question frameworks, and then generated questions. Participants who were assigned to the MDQI After Reading condition studied the documents, viewed the instructional video and question frameworks, and then generated questions. Participants in the control condition carefully read and studied the documents and then generated questions. Participants in all three conditions had the opportunity to take notes on a paper copy of the notetaking matrix while studying the documents online within the survey system. After studying the documents, the participants advanced to the next screen and were unable to return to the documents, but participants were instructed that they could use their notes from the documents to assist them with question generation. All three
conditions generated questions that will help them to understand, evaluate, and synthesize the information from the multiple documents with the same prompt, generating up to five questions. On the following screen, participants in all three conditions responded to the following prompt, “Select the question that BEST helps you to understand, evaluate, and synthesize the information from the multiple documents you just read.” Next, the participants answered the essay prompt in an essay-sized text box. Finally, participants answered a 10-item document knowledge measure. Sessions were two hours long and time was recorded throughout each section of the study. There were no time restrictions throughout the study to allow participants to work through the materials at their own pace.

Table 3

Order of Study Activities by Condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Study Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDQI Before</td>
<td>MDQI Instruction Documents Questions Questions Response Knowledge Posttest</td>
</tr>
<tr>
<td>MDQI After</td>
<td>MDQI Documents Instruction Questions Questions Response Knowledge Posttest</td>
</tr>
<tr>
<td>Control</td>
<td>Prompt to Study Documents Questions Questions Response Knowledge Posttest</td>
</tr>
</tbody>
</table>
**Question coding rubric.** The question coding rubric used in the studies by Cameron and colleagues (2017 & 2018) was used to code each of the questions that the participants generated. The rubric, shown in Table 4, consists of five categories aimed at three levels of stances towards text comprehension (Croninger, Li, Cameron, & Murphy, 2018). The first level consists of questions aimed at fact-finding, an efferent stance, and questions aimed at emotional connections to the documents, an expressive stance. Fact questions can be answered by simply recalling information directly from a single document and are aimed at basic comprehension of the text. For example, “What is imperialism?” is a question that requires a definition of a term that is central to comprehending the documents. Affect questions are aimed at understanding the human experience or human emotions relevant to the documents. For example, “What were the thoughts of the US occupation of the Philippines for the Filipino people?” The second level, the critical-analytic stance, consists of questions aimed at multiple document thinking heuristics, but only refer to a single document. Contextualization questions are aimed at understanding a document in relationship to the time period and what was taking place in that point in time to provide context to the document. For example, “What is the significance of this political cartoon to this time period?” Sourcing questions are related to the author, who the author is, what their purpose is in creating the document, and how that impacts the credibility of the document. Lastly, the third level, the integration stance, consists of questions aimed at multiple document integration, corroborating arguments, viewpoints, concepts, and ideas from multiple sources. These questions require the use of multiple documents in answering the question. This category is what has been referred to throughout the manuscript as the MDQ. A “No Code” category code was reserved for questions that were unrelated to the documents’ topic or questions where it could not be inferred what the participant is asking. For example, “How does the Boxer Rebellion affect us
today?” received a “No Code” as the Boxer Rebellion was unrelated to the documents. Another example of a “No Code,” “How do people choose to make a difference?” It was unclear as to what the participant was asking and how it was related to the documents. Two raters coded each of the questions and reconciled any discrepancies (intrerrater reliability = 78%).
<table>
<thead>
<tr>
<th>Rubric Level</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-Level</td>
<td><strong>Efferent</strong></td>
<td></td>
</tr>
<tr>
<td>Questions</td>
<td>Fact</td>
<td>“What were the benefits to the U.S. of occupation in the Philippines?”</td>
</tr>
<tr>
<td></td>
<td><strong>Expressive</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Affect</td>
<td>“How does US Imperialism affect hard-working, normal families in the Philippines?”</td>
</tr>
<tr>
<td>High-Level</td>
<td><strong>Critical Analytic</strong></td>
<td></td>
</tr>
<tr>
<td>Questions</td>
<td>Contextualization</td>
<td>“What were the actual conditions of the Philippines before America wanted to imperialize?”</td>
</tr>
<tr>
<td></td>
<td><strong>Sourcing</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A question of authorship or purpose</td>
<td>“How reliable is the author of this source?”</td>
</tr>
<tr>
<td>“Multiple</td>
<td><strong>Integration</strong></td>
<td></td>
</tr>
<tr>
<td>Document</td>
<td>Corroboration</td>
<td></td>
</tr>
<tr>
<td>Questions</td>
<td>A question indicating a desire for additional information that contained in another document.</td>
<td>“Were the views expressed in the various documents consistent with popular views of Americans during this time period?”</td>
</tr>
</tbody>
</table>
**Document integration essay coding rubric.** Rukavina and Daneman (1996) developed the following rubric and has been used in multiple document learning research by adapting the rubric to the material content (Bråten et al., 2013; Cameron et al., 2018). The rubric is in Table 5. The author and a member of the research team used this rubric to code each of the responses learners generated. The essay prompt described above directly relates to the coding rubric to prompt for the description and textual support of multiple perspectives from the document set, as was done in previous research (Bråten et al., 2013; Rukavina & Daneman, 1996). A (0) describes no perspective or provides irrelevant or inaccurate information. In the example below, the participant states that the government was going to fund education in the Philippines. While education was discussed, government funding of such education is not discussed in any of the documents. It also describes multiple “territories” fighting over the land of the Philippines, which is again, inaccurate. The conflict was between the United States and the Philippines and United States occupation of the territory. It is also unclear as to what perspective is being described with these fundamental inaccuracies in the comprehension of the document set, receiving a 0 code. A (1) only describes one perspective with no explanation or reason. The example below illustrates a single perspective supporting imperialism due to trade, but does not provide further explanation or textual support. A (2) describes one perspective and provides an explanation of the perspective. A (3) describes both perspectives without providing an explanation for either. In the example the thesis of box texts is stated, but there is no further explanation or textual support to explain the authors stances. A (4) describes both perspectives and provides an explanation for one of perspective. Below, the participant describes the thesis of both texts, but only supports Beveridge’s point of view with evidence related to strategic trade
with China. A (5) describes both perspectives and providing explanation of both perspectives. In the example below, both perspectives are clearly defined with textual support for both perspectives in the description of increased trade from Bryan and in the explanation of foundational principles in American policies from Beveridge.

Similar to the Bråten et al. (2013) study, coding was conducted in a two-step process. First, raters determined how many perspectives were included in the response. Second, raters determined if each perspective was supported by an explanation and evidence from the documents. If there were fundamental misconceptions in the responses or if the responses were entirely comprised of information unrelated to the document set, the response received a 0 code.
### Table 5

**Response Coding Rubric**

<table>
<thead>
<tr>
<th>Code</th>
<th>Category</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Not Accurate</td>
<td>Does not describe a perspective or provides irrelevant information.</td>
<td>“The good thing to come out of it if America gets the territory is that the Filipinos will have a great education because it will be funded by the United States government. But the problem is, is that other territories want this land and we are trying to fight for it.”</td>
</tr>
<tr>
<td>1</td>
<td>Single, No Support</td>
<td>Only describes one perspective with no explanation or reason.</td>
<td>“The conflict between the Philippines and American imperialism was the issue of trade. Trade conflict led to undesired circumstances that changed our history. With Asia being the largest trade force, the Philippines were in a good position.”</td>
</tr>
<tr>
<td>2</td>
<td>Single, Supported</td>
<td>Describes one perspective and explains the perspective.</td>
<td>“The Americans are trying to remain in power of the Philippines. In the text, it described how giving the Philippines education would be a bad thing. This is because giving them education would give them the power and knowledge to revolt against us and possibly gain power over us. We do not want to give them this power and we need to stay in control of them to protect our country and our people.”</td>
</tr>
<tr>
<td>3</td>
<td>Both, No Support</td>
<td>Describes both perspectives without explaining either.</td>
<td>“Albert Beveridge gave his long winded speech last week addressing our president and making strong arguments as to why we are not conquering the last open island in the sea. I actually heard another speech by William Jennings Bryan the other day opposing everything that Beveridge had said. He believes that it is unjust to turn all Filipinos to imperialism out of fear that they could then rise against us with their newfound knowledge and skills.”</td>
</tr>
<tr>
<td>4</td>
<td>Both, Single Support</td>
<td>Describes both perspectives and explains one perspective.</td>
<td>“The first article was titled William Jennings Bryan Opposes US Occupation of the Philippines. This author feels as though we should not be involved with the Philippines because we are only giving them the capability to get stronger and harm the United States...The second article is entitled Albert Beveridge Defends US, this authors viewpoint is very different then the first but they do agree on one concept the Philippines needs help. This author thinks that we should be in control and have a presence of the Philippines because of its access to China and considering how weak they are they could have easily been taken over by China. We also are able to use them as an economic source that will only aid our country because we will no longer need to import as many goods now that we have direct access to them in the Philippines.”</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>5</td>
<td>Both, Supported</td>
<td>Describes both perspectives and explains both perspectives.</td>
<td>“This argument is summarized by Albert Beveridge's argument for US Imperialism. Beveridge believed that it would be essential for the future of the nation to keep hold of the territory for the sole reason that it would be to great economic loss. Beveridge argued that, no matter the humanitarian aspect of keeping the colony, we should keep the Philippines so as to have a gateway into the commerce of Asia. He supported this argument with the fact that European imperialism...This is well characterized by William Jennings Bryant's standpoint on the matter. Principally, Jennings Bryant believed that the United States must let go of the Philippines for fear of our past coming back to bite us. Essentially, he disproved the argument for imperialism by stating that 'though the people of the Philippines may receive better education, they will also receive the ability to read the American Declaration and Constitution.' It is implied that Jennings Bryant feared a revolt against the United States, just as the adolescent nation had risen against imperialist Britain not long before hand. Therefore, the United States must leave the Philippines to themselves for the sake of avoiding hypocrisy.”</td>
</tr>
</tbody>
</table>
Chapter 4

Results

Preliminary Analyses

Two One-way ANOVAs were conducted to ensure that there were no prior differences between the conditions before proceeding to tests of research hypotheses. The first ANOVA compared differences across the three conditions on prior knowledge and reading comprehension served as the dependent variable in the second ANOVA. There were no significant differences between conditions for either prior knowledge \([F(2, 106) = 3.05, p = .051, \eta_p^2 = .05]\) or reading comprehension ability \([F(2, 106) = 1.30, p = .28, \eta_p^2 = .02]\). It is important to note that these results may be affected by the low reliability estimates of both the prior knowledge and reading comprehension measures. The descriptives in Table 5, however, do not seem to indicate any significant deviations by condition which lends support to the findings above. Still, due to this poor reliability, these variables are not considered in further analyses. While the MDQI is expected to improve the quality of participants’ self-generated questions, this instruction is not expected to effect the number of questions that are generated. A third One-way ANOVA, with the number of questions generated after reading, supported this assumption, \([F(2, 106) = .26, p = .77, \eta_p^2 = .01]\). This finding is important because it shows that any effects of condition found in this study cannot be attributed to a more general effect on the tendency to generate questions.

Another important assumption is that condition results cannot be attributed to effort as measured by time spent on the experimental tasks between conditions. A fourth One-way ANOVA supported this assumption, \([F(2, 106) = .11, p = .90, \eta_p^2 = .002]\). All ANOVAs met the assumptions of normality, independence, and homogeneity. All chi-square tests met the assumptions of independence and level of measurement.
Table 5

*Means and Standard Deviations for the Prior Knowledge, Reading Comprehension, & Questions*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Number of Participants</th>
<th>Prior Knowledge</th>
<th>Reading Comprehension</th>
<th>Number of Questions Generated</th>
<th>Time (in Sec.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDQI-AR</td>
<td>36</td>
<td>3.86 (1.79)</td>
<td>3.22 (1.33)</td>
<td>4.03 (1.27)</td>
<td>2126.28</td>
</tr>
<tr>
<td>MDQI-BR</td>
<td>37</td>
<td>4.54 (1.32)</td>
<td>2.64 (1.57)</td>
<td>3.95 (1.37)</td>
<td>2032.16</td>
</tr>
<tr>
<td>Control</td>
<td>36</td>
<td>4.81 (1.86)</td>
<td>2.86 (1.68)</td>
<td>3.81 (1.31)</td>
<td>2040.86</td>
</tr>
</tbody>
</table>

**Question Generation**

After reading, all participants were prompted to generate questions that would help them to understand, evaluate, and synthesize the information from the multiple documents they read. Participants could generate up to five questions. Hypothesis one states that there will be a significant association between condition and the types of questions students generated. That is, participants in both question instruction conditions would generate a greater number of MDQs (the corroboration category below) and fewer lower-level questions than the control condition. It is also expected that participants in the MDQI-BR condition would generate more MDQs and fewer lower level questions than MDQI-AR. A chi-square indicates there is a significant association between condition and question type \[\chi^2(8, N = 109) = 50.46, p < .001, V = .26\]. This is a moderately strong effect size. Table 6 shows the frequencies and percentages of questions generated by condition. A standardized residual represents deviation from what is expected within an individual cell and is significant at one standard deviation or greater in either
direction, positive or negative. The expected value of each cell is calculated from the column
total, divided by the overall total, and multiplied by the row total number. This calculation tests
the assumption that there is no relationship between the condition and question type for each cell.
The standardized residuals in Table 6 indicate that participants in the MDQI-AR condition
generated significantly fewer fact-level questions than expected, while participants in the control
condition generated significantly more fact-level questions than expected. Additionally, the
standardized residuals indicate that the control condition also generated significantly fewer
contextualization questions and fewer corroboration questions (i.e., MDQs) than the expected
value. Another interesting finding is that the control condition participants also generated
significantly more affect questions than the expected value.

These findings replicate the results of Cameron et al. (2016, 2018) and are partially
consistent with Hypothesis one. Those in the both the MDQI-BR and MDQI-AR conditions
generated more corroboration-level questions and fewer fact-level questions than the control
condition, but MDQI-BR did not generate more MDQs than MDQI-AR as predicted. In fact,
MDQI-BR participants generated 4% fewer corroboration questions and 5% more fact questions
than MDQI-AR. The MDQI-BR condition, however, was expected to generate more
corroboration questions and fewer fact questions than the MDQI-AR condition. Although this is
not significant effect, this trend indicates that the timing of question instruction may influence
the types of questions learners generate. It is possible that the task of generating MDQs while
reading became a distraction for MDQI-BR participants by providing the learner with two
competing tasks, thinking about question generation and comprehension, may have impacted
question generation. MDQI-AR, on the other hand, provided participants with question
generation as a reflective activity following first pass comprehension of the text.
Table 6

*Question Type Frequencies, Proportions of Question Types, and Standardized Residuals*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Statistic</th>
<th>No Code</th>
<th>Fact</th>
<th>Affect</th>
<th>Context.</th>
<th>Sourcing</th>
<th>Corroboration</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDQI-AR</td>
<td>%</td>
<td>14%</td>
<td>15%</td>
<td>9%</td>
<td>13%</td>
<td>15%</td>
<td>34%</td>
</tr>
<tr>
<td>MDQI-BR</td>
<td>%</td>
<td>20%</td>
<td>20%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>30%</td>
</tr>
<tr>
<td>Control</td>
<td>%</td>
<td>10%</td>
<td>39%</td>
<td>24%</td>
<td>4%</td>
<td>5%</td>
<td>18%</td>
</tr>
<tr>
<td>MDQI-AR</td>
<td>Freq</td>
<td>20</td>
<td>21</td>
<td>14</td>
<td>19</td>
<td>21</td>
<td>49</td>
</tr>
<tr>
<td>MDQI-BR</td>
<td>Freq</td>
<td>29</td>
<td>29</td>
<td>14</td>
<td>15</td>
<td>14</td>
<td>45</td>
</tr>
<tr>
<td>Control</td>
<td>Freq</td>
<td>14</td>
<td>53</td>
<td>33</td>
<td>5</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>MDQI-AR</td>
<td>Res</td>
<td>-</td>
<td>-2.38</td>
<td>-1.49</td>
<td>1.57</td>
<td>1.77</td>
<td>1.33</td>
</tr>
<tr>
<td>MDQI-BR</td>
<td>Res</td>
<td>-</td>
<td>-0.71</td>
<td>-1.27</td>
<td>0.7</td>
<td>0.14</td>
<td>1.09</td>
</tr>
<tr>
<td>Control</td>
<td>Res</td>
<td>-</td>
<td>3.08</td>
<td>2.73</td>
<td>-2.25</td>
<td>-1.91</td>
<td>-2.40</td>
</tr>
</tbody>
</table>

Note: No Code were excluded from the analysis.

**Question Selection**

After generating questions, participants were instructed to select the best question that helps them to understand, evaluate, and synthesize the information from the multiple documents they read. Table 7 displays the number and proportion of participants who generated at least one MDQ by condition. It also shows the number and proportion of participants who selected an MDQ from the questions they generated by condition. Hypothesis two predicted that there will be a significant association between condition and the selection of a multiple document question. A chi-square test indicated there is no significant association between the condition and selection of corroboration level questions \( \chi^2(2, N = 109) = 5.16, p = .08, V = .22 \). Table 7 also shows that 67% of the participants in MDQI-AR who generated an MDQ selected an MDQ, 60% of MDQI-BR participants, and 64% of control participants. Again, the MDQI-BR condition was
expected to outperform the MDQI-AR and control conditions, but actually performed worse than both conditions, signifying that timing of instruction may also negatively influence learners’ abilities to recognize an MDQ. This pattern shows that, overall, question instruction did not support the learner’s ability to recognize an MDQ, contrary to Hypothesis two. Question instruction may be improved by helping learners recognize the elements of an MDQ and providing practice with feedback to select the best types of questions to foster multiple document learning.

Table 7

*Question Generation and Selection Frequencies, Proportions, and Standardized Residuals*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Statistic</th>
<th>Generated MDQ?</th>
<th>Total</th>
<th>Selected MDQ?</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>MDQI-AR</td>
<td>%</td>
<td>67%</td>
<td>33%</td>
<td>100%</td>
<td>44%</td>
</tr>
<tr>
<td>MDQI-BR</td>
<td>%</td>
<td>54%</td>
<td>46%</td>
<td>100%</td>
<td>32%</td>
</tr>
<tr>
<td>Control</td>
<td>%</td>
<td>31%</td>
<td>69%</td>
<td>100%</td>
<td>19%</td>
</tr>
<tr>
<td>MDQI-AR</td>
<td>Freq</td>
<td>24</td>
<td>12</td>
<td>36</td>
<td>16</td>
</tr>
<tr>
<td>MDQI-BR</td>
<td>Freq</td>
<td>20</td>
<td>17</td>
<td>37</td>
<td>12</td>
</tr>
<tr>
<td>Control</td>
<td>Freq</td>
<td>11</td>
<td>25</td>
<td>36</td>
<td>7</td>
</tr>
<tr>
<td>MDQI-AR</td>
<td>Res</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.31</td>
</tr>
<tr>
<td>MDQI-BR</td>
<td>Res</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.03</td>
</tr>
<tr>
<td>Control</td>
<td>Res</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-1.34</td>
</tr>
</tbody>
</table>

*Condition Effects on Learning*

Following question generation and selection, all participants wrote an essay response in which they explained the historical events that were described in the documents. The essay
responses were scored for both the number of perspectives included in the response and the textual evidence included in the response. Hypothesis three predicted there would be a significant association between condition and multiple document integration. A 3 x 6 chi-square analysis indicated no significant association between condition and essay quality \( \chi^2(10, N = 109) = 17.57, p = .06, V = .40 \). Hypothesis three was not supported, participants in the MDQI conditions did not perform significantly better on scores of multiple document integration than the control condition as seen in Table 8. It should also be noted that 43% of the MDQI-BR condition participants’ responses were inaccurate compared to only 22% of control condition participants’ responses and 28% of MDQI-AR participants’ responses. Additionally, only 35% of MDQI-BR participants included multiple perspectives in their responses compared to 61% of MDQI-AR participants and 59% of Control participants. The notetaking matrix artifacts provide some explanation for this pattern in that 41% of MDQI-BR participants failed to take notes on any of the six documents compared to only 14% of MDQI-AR participants and 22% of Control participants. It is possible that by providing the learners with MDQI before reading may have distracted them or increased cognitive load during the reading task, negatively impacting the learning process.
Table 8

*Response Type Frequencies, Proportions, and Standardized Residuals by Condition*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Statistic</th>
<th>Inaccurate</th>
<th>Single</th>
<th>Single + Support</th>
<th>Both</th>
<th>Both + Single Support</th>
<th>Both + Both Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDQI-AR</td>
<td>%</td>
<td>28%</td>
<td>8%</td>
<td>3%</td>
<td>25%</td>
<td>8%</td>
<td>28%</td>
</tr>
<tr>
<td>MDQI-BR</td>
<td>%</td>
<td>43%</td>
<td>16%</td>
<td>5%</td>
<td>5%</td>
<td>11%</td>
<td>19%</td>
</tr>
<tr>
<td>Control</td>
<td>%</td>
<td>22%</td>
<td>3%</td>
<td>16%</td>
<td>22%</td>
<td>6%</td>
<td>31%</td>
</tr>
<tr>
<td>MDQI-AR</td>
<td>Freq</td>
<td>10</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>MDQI-BR</td>
<td>Freq</td>
<td>16</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Control</td>
<td>Freq</td>
<td>8</td>
<td>1</td>
<td>6</td>
<td>8</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>MDQI-AR</td>
<td>Res</td>
<td>-.37</td>
<td>-.16</td>
<td>-1.14</td>
<td>1.09</td>
<td>.02</td>
<td>.25</td>
</tr>
<tr>
<td>MDQI-BR</td>
<td>Res</td>
<td>1.32</td>
<td>1.42</td>
<td>-.61</td>
<td>-1.75</td>
<td>.54</td>
<td>-.81</td>
</tr>
<tr>
<td>Control</td>
<td>Res</td>
<td>-.96</td>
<td>-1.26</td>
<td>1.76</td>
<td>.69</td>
<td>-.56</td>
<td>.58</td>
</tr>
</tbody>
</table>

After responding to the essay prompt, participants completed a 10-item multiple-choice posttest to assess their knowledge of the relevant document information from each document.

Five items assessed content from the text documents, and five items assessed content from the visual documents. Hypothesis four predicted that MDQI participants would score higher on a measure of document knowledge than Control condition participants, with MDQI-BR participants scoring the highest on document knowledge. A One-way ANOVA indicated that there are no significant differences between conditions on document posttest knowledge \(F(2, 106) = 1.30, p = .28, \eta^2 = .02\). These results, however, may be affected by the low reliability of the document knowledge posttest scores. An item-analysis indicated that the measure included items a good range of difficulty levels, with the exception of item 9, which no participant
answered correctly. Items 2, 3, 8, and 10 had poor item-total correlations (i.e., below .2) which could also be contributing to the poor reliability. Item 2 was a true-false item with a p-value of .76, but poor item-total correlation, which could be an indication that there was high number of guessing on this item. Items 3, 8, and 10 are items related to the text documents which is also reflected in the lower average text document scores. The p-values of items 3, 8, and 10 were .35, .41, and .35 respectively, the most difficult items in the measure. The descriptive statistics in Table 9 lend further support to this analysis by showing virtually no differences between conditions.

Table 9

<table>
<thead>
<tr>
<th>Condition</th>
<th>Posttest Total</th>
<th>Posttest Text Document Items</th>
<th>Posttest Visual Document Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDQI-AR</td>
<td>5.86</td>
<td>2.36</td>
<td>3.50</td>
</tr>
<tr>
<td></td>
<td>(1.94)</td>
<td>(1.19)</td>
<td>(1.21)</td>
</tr>
<tr>
<td>MDQI-BR</td>
<td>5.91</td>
<td>2.43</td>
<td>3.48</td>
</tr>
<tr>
<td></td>
<td>(1.97)</td>
<td>(1.31)</td>
<td>(1.24)</td>
</tr>
<tr>
<td>Control</td>
<td>5.57</td>
<td>2.08</td>
<td>3.49</td>
</tr>
<tr>
<td></td>
<td>(1.97)</td>
<td>(1.21)</td>
<td>(1.19)</td>
</tr>
</tbody>
</table>

**MDQ Generation Effects on Learning**

Participants were divided into two groups, those who generated an MDQ and those who did not, to test the effects of the self-generated MDQ on multiple document learning. Table 7 shows the distribution of participants who did and did not generate at least one MDQ by condition. The majority of participants in the MDQI conditions generated at least one MDQ,
MDQI-AR (67%) and MDQI-BR (54%), and only 31% of the control condition participants generated at least one MDQ. A 2 x 6 chi-square analysis tested for an association between MDQ generation and the number of document perspectives and evidential support included the essay response [$\chi^2(5, N = 109) = 11.05, p = .05, V = .32$]. This association shows that learners who tend to generate MDQs also tend to report a more integrated representation of the content within the document set. In fact, an inspection of the score distributions shown in Table 10 indicates that 61% of the participants who failed to generate an MDQ also failed to report an understanding of the documents that included both perspectives. For those who did generate MDQs, by contrast, 64% did include both perspectives in their explanation of the events.

Analyses of time and the notetaking matrix artifacts provide further evidence that the possibility that this outcome is not attributed solely to effort. There is not a significant correlation between time and performance on the response measure [$r (107) = .08, p = .41$]. Also, there is not a significant correlation between the number of documents participants included in the notetaking matrix and response [$r (107) = .10, p = .30$] nor is there a significant correlation between notetaking matrix word count and response [$r (107) = -.02, p = .84$].
Table 10

MDQ Response Type Frequencies, Proportions, and Standardized Residuals

<table>
<thead>
<tr>
<th>Generated MDQ</th>
<th>Statistic</th>
<th>Inaccurate</th>
<th>Single</th>
<th>Single + Support</th>
<th>Both</th>
<th>Both + Single Support</th>
<th>Both + Both Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>%</td>
<td>39%</td>
<td>9%</td>
<td>13%</td>
<td>19%</td>
<td>4%</td>
<td>17%</td>
</tr>
<tr>
<td>Yes</td>
<td>%</td>
<td>24%</td>
<td>9%</td>
<td>3%</td>
<td>16%</td>
<td>13%</td>
<td>35%</td>
</tr>
<tr>
<td>No</td>
<td>Freq</td>
<td>21</td>
<td>5</td>
<td>7</td>
<td>10</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Yes</td>
<td>Freq</td>
<td>13</td>
<td>5</td>
<td>2</td>
<td>9</td>
<td>7</td>
<td>19</td>
</tr>
<tr>
<td>No</td>
<td>Res</td>
<td>1.01</td>
<td>.02</td>
<td>2.29</td>
<td>.19</td>
<td>-1.17</td>
<td>-1.40</td>
</tr>
<tr>
<td>Yes</td>
<td>Res</td>
<td>-1.00</td>
<td>-.02</td>
<td>-1.19</td>
<td>-.19</td>
<td>1.15</td>
<td>1.30</td>
</tr>
</tbody>
</table>

Hypothesis six predicts that there will be a significant difference between those who generate an MDQ and those who do not generate an MDQ on document knowledge. An independent t-test indicated significant differences between those who did and did not generate an MDQ on document posttest knowledge \( t(107) = 2.07, p = .04, d = .39 \). Results support Hypothesis six; participants who generated an MDQ had higher posttest scores on average than participants who did not generate an MDQ. Table 11 shows that the descriptive statistics provide further support to these findings as they indicate almost a full point difference in average posttest score between MDQ and No-MDQ groups. These results, however, may be affected by the low reliability of the document knowledge posttest scores. Again, item analyses revealed some concerns with the document knowledge posttest score reliability, so interpretations must be considered with this in mind.
Table 11

*MDQ Means and Standard Deviations for the Document Knowledge Posttest*

<table>
<thead>
<tr>
<th>Generated MDQ</th>
<th>Number of Participants</th>
<th>Posttest Total</th>
<th>Posttest Text Document Items</th>
<th>Posttest Visual Document Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>55</td>
<td>6.29</td>
<td>2.56</td>
<td>3.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.97)</td>
<td>(1.27)</td>
<td>(1.78)</td>
</tr>
<tr>
<td>No</td>
<td>54</td>
<td>5.52</td>
<td>2.29</td>
<td>3.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.92)</td>
<td>(1.35)</td>
<td>(1.27)</td>
</tr>
</tbody>
</table>

**Summary of Results**

MDQI does appear to support learner’s abilities to generate MDQs, but did not support learner’s abilities to recognize and select an MDQ if they generated one. There were no condition effects on multiple document learning on either the multiple document integration response measure or the document knowledge posttest. Item analyses revealed some concerns with the document knowledge posttest score reliability, so interpretations must be considered with this in mind. Importantly, MDQI-BR was expected to outperform MDQI-AR, but actually performed worse than control condition participants—and also MDQI-AR study participants—on MDQ selection and on the multiple document integration response measure. Providing MDQI before reading may have been distracting to the participants and again, may have increased cognitive load during the reading task.

To directly test the effects of MDQ generation on multiple document learning, participants were separated into those who did or those who did not generate an MDQ. The analyses and patterns of finding indicate that MDQ generation has a moderate effect on multiple document integration and knowledge, but further instructional support and more engaged deep-
processing of the texts in notetaking and recursive second or third pass reading could lead to improved posttest scores and integrated responses by MDQ participants.
Chapter 5

Discussion

This study tests the effects of Multiple Document Question Instruction (MDQI; Cameron et al., 2017), a brief strategy instruction intervention that teaches learners how to generate high-quality multiple document questions (MDQs). This instruction is intended to increase the number of MDQs learners generate, improve the ability to recognize an MDQ, and increase attention to multiple texts to improve multiple document learning. It also tests the effects of MDQ generation on learning. This chapter situates the findings of this study within the broader theoretical and practical implications and discusses limitations and future directions of the research.

Summary of Findings

Question generation. After reading the documents in the study, all participants were prompted to generate questions that would help them to understand, evaluate, and synthesize the information from the multiple documents they read. There was a significant association between condition and the types of questions participants generated. This significance can be attributed to control participants generating significantly fewer MDQs and more fact questions than the expected value, and the MDQI-AR participants generated significantly more MDQs and fewer fact level questions than the expected value. These findings are consistent with previous research (Cameron et al. 2017, 2018). Contrary to expectations, the MDQI-BR condition did not outperform both the control and MDQI-AR conditions. While the MDQI-BR condition did generate a greater number of MDQs and fewer fact questions than the control condition, the MDQI-BR condition generated fewer MDQs and more fact questions than MDQI-AR. This finding suggests that the timing of question instruction could negatively impact the ability to generate MDQs.
**Question selection.** After generating questions, participants were instructed to select the best question that helps them to understand, evaluate, and synthesize the information from the multiple documents they read. There was no significant association between the condition and selection of MDQ (corroboration level) questions. Less than half of the MDQI-AR participants and a third of the MDQI-BR participants who generated an MDQ were able to select an MDQ. Worse yet, approximately 20% of the participants in each condition who generated an MDQ, failed to select an MDQ. By comparison, only 19% of control participants who generated an MDQ selected an MDQ, with 12% of control participants who did generate an MDQ failing to select an MDQ. These findings suggest that although MDQI increases the number of MDQs participants generate they still struggle to identify these types of questions without further instructional support.

**Condition effects on learning.** Multiple document learning was measured in two different ways in this study. First, the essay response measured the recall of information from the documents by the inclusion of multiple perspectives and the evidence from the documents. Hypothesis three was not supported, and there was not a significant association between condition and multiple document integration. Specifically, participants in the MDQI conditions did not perform significantly better on scores of multiple document integration than the control condition. In fact, the MDQI-BR condition participants performed worse than both MDQI-AR condition and control participants. Most concerning is that almost half of the MDQI-BR participants produced inaccurate responses.

Second, the document knowledge posttest score measured the comprehension of document information. Hypothesis four predicted that there would be a significant difference between conditions on document knowledge. MDQI participants are expected to score higher on
a measure of document knowledge than control condition participants, with MDQI-BR participants scoring the highest on document knowledge. There were no significant differences between conditions on document knowledge. Overall, there were no significant condition effects on learning, but the timing of MDQI before reading could actually negatively impact learning. This is limited by the performance of the measure as further discussed in the limitation section.

**MDQ generation effects on learning.** Hypothesis five and six were intended to test the direct effects of MDQ generation on learning, regardless of condition assignment. Participants were divided into two groups to test the effects of the self-generated MDQ on multiple document learning: those who generated an MDQ and those who did not. There was a significant association between MDQ generation, and the number of document perspectives and evidential support included in the essay response. Results also supported hypothesis six; participants who generated an MDQ had higher posttest scores on average than participants who did not generate an MDQ. Again, these findings are limited to the performance of the measure as further discussed in the limitation section below.

**Conclusions.** This study's results provide further evidence for the viability of question instruction to improve the quality of questions that learners generate when learning from multiple documents. The literature demonstrates that students often struggle to generate quality questions that go beyond factual recall (Chin & Osborne, 2008; Cameron et al., 2017; Cameron et al., 2018). Participants in both MDQI conditions generated a higher number of corroboration questions, questions aimed at multiple document comprehension, than their peers in the control condition. Question instruction is necessary to improve the quality of questions learners generate, but the timing of question instruction is important in maximizing the potential of students to generate questions. The results of this study, that is, suggest that providing learners
with question instruction after they read and study the documents is more desirable than providing this instruction before reading.

Findings also indicate that MDQI did not support the learner’s ability to recognize an MDQ. Learners may need more examples and time to practice identifying MDQs with feedback to support their understanding of what an MDQ is and how to recognize an MDQ. In addition, timing of the question instruction may have also negatively impacted MDQ selection. There are some important differences in question selection between the MDQI-BR condition participants and MDQI-AR condition participants, especially in the context of question generation. Not only did the MDQI-BR group generate 13% fewer MDQs than the MDQI-AR group, but they also selected 11% fewer MDQs than the MDQI-AR participants.

Along with the results of question generation and selection, question training did not have a greater impact on multiple document learning when experimental participants received MDQI before reading. It is possible that the task of generating questions while reading hindered comprehension while reading. There appears to be a compounded effect of the order of instruction for the experimental conditions that carries from question generation through the learning measures in this study.

More encouraging though, are the findings that show learners who tend to generate MDQs produce responses with multiple perspectives supported with textual evidence. MDQ generation appears to direct attention to multiple document information in a meaningful way that also improves the memory of the documents to support the learner’s response. This finding is significant as it demonstrates the potential for multiple document learning by directly using the self-generated question strategy. This is ultimately the purpose of MDQI.
Implications of Findings

Self-directed multiple document learning. Much of the previous research done in multiple document learning has focused on the effects of researcher-directed tasks to influence the selection and integration of multiple document information. While this research has provided valuable information regarding how learners interact with documents and task specifications, it does not address the need for self-sustaining instructional supports for learners. The purpose of this research is to address this need by training students on how to generate questions that can be used to learn from multiple documents. It is also possible that because self-generated questioning is a learner-directed strategy and not a researcher-directed task, learners may be able to transfer this strategy to future multiple document tasks upon further use and training. This is a powerful tool for a life-long learner inside and outside of the classroom. By providing learners with tools to employ throughout their formal and informal learning processes, we increase the robustness of our educational intervention.

Questions as tools for learning. Prior research has demonstrated that learners who use questions that aimed at multiple document comprehension are more likely to recall more information from multiple documents than learners who rely on factual, single-document question types (Cameron et al., 2018; Taboada & Guthrie, 2006). These results provide evidence that the question generation strategy can be used to support learning from multiple documents.

Strategy use is part of the learning processes and not a final task product. The MD-TRACE model depicts the multiple document learning process as being self-regulated and recursive. The question strategy supports the process of multiple document learning as questioning helps to direct attention and selection of multiple document information (Taboada & Guthrie, 2006; Wiley & Voss, 1999), regulate learning by allowing the learner to monitor
comprehension (Graesser & Lehman, 2011), and foster connections within and between documents for integration. While a task can be designed to foster these processes, self-generated questions are flexible enough to change with the individual learner’s needs throughout the learning process.

**Timing matters.** Additionally, the timing of strategy instruction does impact the learning process. The MDQI-BR condition was expected to perform better than both control and the MDQI-AR condition on the basis that providing learners with the knowledge of MDQs prior to reading would assist in reading comprehension. Our findings suggest that providing this instruction before reading actually hindered learners in their ability to learn from the multiple documents. Instead of using question generation as a comprehension monitoring tool, it became a distracting task that took away from understanding the documents.

This finding is significant in that instructors must take caution in timing and implementation of strategy instruction. By providing learners with MDQI after reading, learners were able to use the strategy to reflect on what they understood and what they did not. In an authentic multiple documents process, a question may prompt the learner to take additional passes through the documents to answer that question and generate more questions. This is the recursive nature of multiple document learning that is described in multiple document learning theory, and may be key to the deeper and more integrated knowledge being sought after.

**Limitations**

The primary limitation of this study is the poor reliability estimates for three of the measures. Despite implementing these measures in the same manner and with a similar population to the previous research, the alpha coefficients for the prior knowledge measure and reading comprehension measure were significantly lower than what that research reported. The
reading comprehension measure consisted of items developed for the Graduate Record Exam (GRE). This measure served as a valid measure of reading comprehension in a recent study (Alexander et al., 2016) and the reliability estimate of the scores were high (Cronbach’s alpha = .79). The reliability estimate of the scores for the current study was much lower (Cronbach’s alpha = .28). However, the mean and standard deviation of the current study ($M = 2.92, SD = 1.55$) were similar to that of the previous sample ($M = 3.21, SD = 1.62$). The average is low and could be an indication that participants struggled to read and comprehend the passages. The difficulty of this measure could have led to guessing on many of the items further contributing to the poor reliability.

The prior knowledge measure was also developed and used in previous research (Cameron et al., 2016; 2018), but the reliability estimate of these scores from the study was low (Cronbach’s alpha = .25). The mean and standard deviation were also a full point lower than previous research ($M = 4.40, SD = 1.71$) suggesting that these participants had a lower level of prior knowledge and guessed on items, which could have impacted the reliability of the scores. Another possible explanation for such poor reliabilities is low engagement on these multiple-choice measures.

The purpose of these measures in this study was to ensure that there were no prior differences between the conditions before proceeding to tests of research hypotheses. The descriptive statistics in Chapter 4 do not seem to indicate significant differences between the conditions and lend further support to the preliminary analyses findings. Due to this poor reliability, however, these variables were not considered in further analyses and did not impact the overall findings of the study.
Another limitation was the performance of the document knowledge posttest items. The document knowledge posttest was developed and implemented for the first time in this study to capture participant’s knowledge gains. The purpose of the measure was to understand which documents learners attended to, what important information they were able to select from single documents, and how documents related to each other to understand the complexities of the topic. The reliability estimate of these scores for this measure was also low (Cronbach’s alpha = .50). This poor reliability and poor item performance prevented insight into understanding the effects of MDQI and MDQ generation on learning in the way it was intended. Overall, items need to be further developed and analyzed to improve the reliability and validity of the measure for future implementation and interpretation.

An additional limitation of this study was that the experimental control restricted the flexible use of the self-generation question strategy. Participants were not able to return to the documents after their initial pass. Participants had one opportunity to study the documents before generating and selecting their questions. This has been a common experimental design in multiple document studies in order to isolate the direct effects of the intervention (e.g., Anmarkrud et al., 2014; Gil et al., 2010a; Le Bigot & Rouet, 2007). The idea has been to find the single task that yields the best multiple document learning results. While this design provides insight in what learners select and integrate given a specific task, it does not provide the entire scope of the multiple documents learning process. The MD-TRACE model itself is task driven, but also suggests the process is recursive, allowing for multiple passes and deeper processing of the documents. The results here indicate that a single pass with MDQI is not sufficient for producing the deeper and more integrated learning we would like to see from learners during these tasks. This is to be expected, however, in that the recursive process was not there by
design. This study does help to determine that question instruction can increase the number of MDQs learners generate which leads to greater attention to multiple document information. This was an important first step in determining the potential for self-generated questions in multiple document learning.

**Future Research**

The current study provides insight into future directions in the development of both MDQI and the self-generated questioning strategy to support multiple document learning. The findings indicate that MDQI does support MDQ generation, and MDQI could be further developed to help the learner understand the different types of questions and when to use each type of question. The results of the question selection analysis reveal that there needs to be additional clarity to what the types of questions are and when to use questions to support multiple document comprehension. Future research could further develop MDQI to provide such clarity.

In order to more completely capture the effects of MDQs on learning, future research could consider allowing participants to work with the documents more recursively. Authentic multiple document learning tasks may require multiple passes through each of the documents, as suggested by the MD-TRACE model (Britt & Rouet, 2011). By allowing participants continual access to reread the documents, participants could more flexibly use the self-generated question strategy throughout the learning process. For example, the participant could read, generate a question, and then return to the reading to find an answer to that question. This process could aid in document comprehension and monitoring (Graesser & Lehman, 2011). Researchers could capture process data while participants work more recursively to capture how participants use the strategy and test for strategy effects on learning.
To test the full effects of this strategy, we must also consider the long-term effects on learning. First, future research should address the effects of the strategy on learning with delayed posttests in addition to the immediate posttest. Delayed posttests would address how the learner's knowledge fundamentally changed as a result of the learning experience with the strategy. Second, future research should address the transfer effects of strategy instruction. One of the primary purposes of this study was to test a learner-directed strategy. The goal is to provide learners with self-directed tools in order to successfully learn from multiple documents throughout their lifetime. This requires that this strategy can be transferred beyond the one-time, experimental setting and even across disciplines.

Conclusion

The purpose of this study was to test the effects of the self-generated question strategy on multiple document learning. The findings of this study support that the Multiple Document Question Instruction continues to promote multiple document question generation. The findings also suggest that the generation of these multiple document questions helps to direct attention to multiple document perspectives and content to support learning. While there is still room for improvement in developing question instruction, these results are promising in supporting learner’s abilities to use the self-generated question strategy to learn from multiple documents.
References


Reading Comprehension Sample Questions. (2014). *GRE Revised General Test: Reading comprehension.*


Appendices

Appendix A- Study Materials & Measures

Implied Informed Consent
Demographics Survey
Reading Comprehension Measure
Prior Knowledge Measure
Question Generation & Selection Directions
Document Integration Measure: Essay Response Prompt
Document Knowledge Posttest
Implied Informed Consent

Title of Project: Learning from Historical Information
Principal Investigator: Chelsea Cameron, Graduate Student
                225 CEDAR Building
                University Park, PA 16802
                cec5144@psu.edu
Advisor: Dr. Peggy Van Meter
                228 CEDAR Building
                University Park, PA 16802
                pnv1@psu.edu

1. Purpose of the Study: The purpose of this research study is to explore how college students study and learn from historical information.
2. Procedures to be followed: You will be asked a series of questions about how you think and then you will be given instructions and asked to generate questions based on history readings.
3. Duration: It will take about two hours to complete the entire procedure.
4. Statement of Confidentiality: Your participation in this research is confidential. The researcher would use ID information only to inform the teacher about extra credit. Once this is completed, the ID would be destroyed or stored separately from the data. In the event of any publication or presentation resulting from the research, no personally identifiable information will be shared because your name is in no way linked to your responses. Your confidentiality will be kept to the degree permitted by the technology being used. No guarantees can be made regarding the interception of data sent via the Internet by any third parties.
5. Right to Ask Questions: Please contact Chelsea Cameron at cec5144@psu.edu with questions or concerns about this study.
6. Payment for participation: Participants will receive 6 extra credit points for their EDPSY 014 course for completion of the survey.
7. Voluntary Participation: Your decision to be in this research is voluntary. You can stop at any time. You do not have to answer any questions you do not want to answer.

You must be 18 years of age or older to take part in this research study.

    o Yes, I am 18 years of age or older.
    o No, I am under the age of 18.

Completion and submission of the survey implies that you have read the information in this form and consent to take part in the research. You may request a copy of this form for your records or future reference.
Demographics Survey

Directions: You will now be asked to complete the following questions regarding your demographic information. Please note this is for research purposes only and your answers are strictly confidential.

1. Indicate your declared or intended major below.

2. Select the number of college history credits you have earned. Including AP credit from high school.
   - 3 or Fewer Credits
   - 4-6 Credits
   - 7-12 Credits
   - More than 12 credits

3. At what level were these history courses in which you received credit for? Check all that apply.
   - Advanced Placement- High School
   - 001-100 level
   - 200 level
   - 300 level
   - 400 level

4. Enter your approximate SAT Verbal section score if applicable.

5. Enter your approximate ACT score if applicable.

6. Enter your current GPA.

7. Please indicate your gender.
   - Male
   - Female
   - Other
8. Is English your first language?
   - Yes
   - No

9. Please indicate your ethnicity.
   - White/Caucasian
   - African American
   - Asian
   - Native American
   - Hispanic
   - Other

10. Please indicate your academic year.
    - Freshman
    - Sophomore
    - Junior
    - Senior
    - Graduate Student
Reading Comprehension Measure

Directions: Read and answer the following 9 items with the best answer choice for each item. *Questions 1 and 2 can be answered using the following paragraph:*

Policymakers must confront the dilemma that fossil fuels continue to be an indispensable source of energy even though burning them produces atmospheric accumulations of carbon dioxide that increase the likelihood of potentially disastrous global climate change. Currently, technology that would capture carbon dioxide emitted by power plants and sequester it harmlessly underground or undersea instead of releasing it into the atmosphere might double the cost of generating electricity. But because sequestration does not affect the cost of electricity transmission and distribution, delivered prices will rise less, by no more than 50 percent. Research into better technologies for capturing carbon dioxide will undoubtedly lead to lowered costs.

1. This passage implies which of the following about the current cost of generating electricity?
   - It is higher than it would be if better technologies for carbon dioxide were available.
   - It is somewhat less than the cost of electricity transmission and distribution.
   - It constitutes at most half of the delivered price of electricity.
   - It is dwelt on by policymakers to the exclusion of other costs associated with electricity delivery.
   - It is not fully recovered by the prices charged directly to electricity consumers.

2. The passage suggests that extensive use of sequestration would, over time, have which of the following consequences?
   - The burning of fossil fuels would eventually cease to produce atmospheric accumulations of carbon dioxide.
   - The proportion of the delivered price of electricity due to generation would rise and then decline.
   - Power plants would consume progressively lower quantities of fossil fuels.
The following question can be answered using the passage below.

Centuries ago, the Maya of Central America produced elaborate, deeply cut carvings in stone. The carvings would have required a cutting tool of hard stone or metal. Iron-ore deposits exist throughout Central America, but apparently the Maya never developed the technology to use them and the metals the Maya are known to have used, copper and gold, would not have been hard enough. Therefore, the Maya must have used stone tools to make the carvings.

3. Which of the following, if true, most seriously weakens the argument?

- In various parts of the world, civilizations that could not make iron from ore fashioned tools out of fragments of iron from meteorites.
- All of the metallic Mayan artifacts that have been found by archeologists are made of metals that are too soft for carving stone.
- The stone out of which these carvings were made is harder than the stone used by the other Central American peoples.
- The technique that the Maya used to smelt gold and some other metals could not have been easily applied to the task of extracting iron from iron ore.
- Archaeologists disagree about how certain stone tools that have been found among Mayan ruins were used.

Directions: Fill in the blank. The following questions have five answer choices. Select the best answer from among these choices.

4. Since she believed him to be both candid and trustworthy, she refused to consider the possibility that his statement had been ____________.

- irrelevant
- facetious
- mistaken
- critical
- insincere
5. It is his dubious distinction to have proved what nobody would think of denying, that Romero at the age of sixty-four writes with all the characteristics of __________.

- maturity
- fiction
- inventiveness
- art
- brilliance

**Directions:** Select the best answer. Questions 6 - 9 can be answered using the following paragraph:

In the 1970’s, two debates engaged many scholars of early United States history. One focused on the status of women, primarily White women. Turning on the so-called golden age theory, which posited that during the eighteenth-century colonial era, American women enjoyed a brief period of high status relative to their English contemporaries and to nineteenth-century American women, this debate pitted scholars who believed women’s lives deteriorated after 1800 against those who thought women’s lives had been no better before 1800. At issue were the causes of women’s subordination: were these causes already in place when the English first settled North America or did they emerge with the rise of nineteenth-century industrial capitalism? The second debate, the so-called origins debate, concerned the emergence of racial slavery in the southern colonies: was slavery the inevitable result of the deep-rooted racial prejudice of early British colonists or did racial prejudice arise only after these planters instituted slave labor? Although these debates are parallel in some respects, key differences distinguished them. Whereas the debate over women’s status revolved around implicit comparisons of colonial women to their counterparts in the antebellum period (1800-1860), thus inviting comment from scholars of both historical periods, the origins debate was primarily confined to a discussion about slavery in colonial America. Second, in contrast to the newness of the debate
over women’s status and its continued currency throughout the early 1980’s, the debate over race and slavery, begun in the 1950’s, had lost some of its urgency with the publication of Morgan’s American Slavery, American Freedom (1975), widely regarded as the last word on the subject. Each debate also assumed a different relationship to the groups whose histories it concerned. In its heyday, the origins debate focused mainly on White attitudes toward Africans rather than on Africans themselves. With few exceptions, such as Wood’s Black Majority (1974) and Mullin’s Flight and Rebellion (1972), which were centrally concerned with enslaved African men, most works pertaining to the origins debate focused on the White architects, mostly male, of racial slavery. In contrast, although women’s historians were interested in the institutions and ideologies contributing to women’s subordination, they were equally concerned with documenting women’s experiences. As in the origins debate, however, early scholarship on colonial women defined its historical constituency narrowly, women’s historians focusing mainly on affluent White women. Over time, however, some initial differences between the approaches taken by scholars in the two fields faded. In the 1980’s, historians of race and slavery in colonial America shifted their attention to enslaved people; interest in African American culture grew, thereby bringing enslaved women more prominently into
view. Historians of early American women moved in similar directions during the decade and began to consider the effect of racial difference on women’s experiences.

6. The passage is primarily concerned with:
   - showing how historians who engaged in a particular debate influenced historians engaged in another debate.
   - explaining why two initially parallel scholarly debates diverged in the 1980's.
   - comparing two scholarly debates and discussing their histories.
   - contrasting the narrow focus of one scholarly debate with the somewhat broader focus of another.
   - evaluating the relative merits of the approaches used by historians engaged in two overlapping scholarly debates.

7. It can be inferred that the author of the passage mentions *American Slavery, American Freedom* primarily in order to
   - substantiate a point about the methodology that came to be prevalent among scholars engaged in the origins debate.
   - cite a major influence on those scholars who claimed that racial prejudice preceded the institution of slavery in colonial America.
   - show that some scholars who were engaged in the origins debate prior to the 1980's were interested in the experiences of enslaved people.
   - identify a reason for a certain difference in the late 1970's between the origins of debate and the debate over American women's status.
   - contrast the kind of work produced by scholars engaged in the origins debate with the kind produced by scholars engaged in the debate over American women's status.

8. The passage suggests which of the following about the women's historians mentioned in the third paragraph?
   - They disputed certain claims regarding the status of eighteenth-century American women relative to women in England during the same period.
   - Their approach to the study of women's subordination had been partly influenced by earlier studies published by some scholars engaged in the origins debate.
   - Their work focused on the experiences of both White and African American women.
   - Their approach resembled the approach taken in studies by Wood and by Mullin in that they were interested in the experiences of people subjected to a system of subordination.
   - To some extent, they concurred with Wood and with Mullin about the origins of racism in colonial America.
9. According to the passage, historical studies of race and slavery in early America that were produced *during* the 1980's *differed* from studies of that subject produced *prior* to the 1980's in that the studies produced *during* the 1980's:

- gave more attention to the experience of enslaved women.
- gave less attention to the cultures of enslaved people.
- were read by more scholars in other fields.
- were more concerned with the institutions and ideologies that perpetuated racial prejudice in postcolonial America.
- made direct comparisons between the subordination of White women and the subordination of African American people.
Prior Knowledge Measure

On the next page, you will be asked a series of questions related to your knowledge of the history of United States Imperialism. Please do your best to answer the next set of questions.

Directions: Please select the best answer for each of the 10 items below.

1. A primary purpose for building the Suez Canal was to
   - encourage Jewish settlement in nearby Palestine.
   - reduce the time needed for travel between the Middle East, Europe and Asia.
   - reduce the time needed for travel between the Atlantic Ocean and the Caribbean Sea.
   - allow Indian merchants to reach the east coast of Africa.

2. Which is an accurate statement about the partitioning of Africa by European imperialist nations during the 1800’s?
   - New nations were based on old tribal boundaries.
   - The cultural and ethnic diversity of the African people was disregarded.
   - The continent was divided equally among the colonial powers.
   - African unity was encouraged.

3. The 19th century term “White Man’s Burden” reflects the idea that
   - Asians and Africans were equal to Europeans.
   - Asians and Africans would be grateful for European help.
   - imperialism was opposed by most Europeans.
   - Europeans had a responsibility to improve the lives of the colonial peoples.

4. The Sepoy Mutiny in India, the Boxer Rebellion in China, and the Islamic Revolution in Iran were similar in that they
   - restored power to the hereditary monarchies.
   - attempted to reject the traditional cultures in these countries.
   - resisted foreign influence in these countries.
   - reestablished the power of religious leaders.
5. Which United States policy is most closely associated with the annexation of Hawaii and the Philippines?
   - Containment
   - Détente
   - Neutrality
   - Imperialism

6. In the past, European nations have conquered other lands, made them into colonies, and controlled their economies. This statement describes a situation that resulted from the
   - industrialization of Europe and the need for raw materials.
   - desire of Europeans to spread communism throughout the world.
   - European belief in human rights for all people.
   - requests of developing nations for modern machines and technology.

7. Which factor is most closely associated with the decision of the United States to declare war on Spain in 1898?
   - Isolationist policy
   - Labor union pressure
   - Yellow journalism
   - Unrestricted submarine warfare

8. The Monroe Doctrine declared that the United States would
   - prevent the establishment of new European colonies anywhere in the world.
   - help colonies in North and South America adopt a democratic form of government.
   - view European interference in the Americas as a threat to the national interest of the United States.
   - prevent other nations from trading with South American nations.

9. Which of the following allowed the US to acquire the Philippines, Puerto Rico, and Guam?
   - Monroe Doctrine
   - 1898 Treaty of Paris
   - Wilson's 14-Point Doctrine
   - 1898 Treaty of Spain
10. Which argument was used to support United States acquisition of overseas possessions in the late 1800’s?

- The United States needed to obtain raw materials and new markets.
- The spread of Marxist ideas had to be stopped because they threatened world peace.
- The United States should be the first world power to build a colonial empire.
- The doctrine of Manifest Destiny had become obsolete.
Question Generation & Selection Directions

Directions: Now generate your own questions in order to synthesize, analyze, and evaluate multiple documents based on the historical documents you examined and read in the previous section. Enter one question in each text box. You may enter as many as 5 questions.

1.

2.

3.

4.

5.

Directions: Select the question that is best suited to synthesize, analyze, and evaluate multiple documents at the same time.

- [Question #1 Generated by Participant from Previous Page]
- [Question #2 Generated by Participant from Previous Page]
- [Question #3 Generated by Participant from Previous Page]
- [Question #4 Generated by Participant from Previous Page]
- [Question #5 Generated by Participant from Previous Page]
Document Integration Measure: Essay Response Prompt

Directions: Write a letter to friend explaining the conflict over the occupation of the Philippines and American imperialism in late 19th and early 20th century. Support your response with evidence from the documents.
Document Knowledge Posttest

Directions: Answer the following questions based on the documents you studied previously.

1. Who wrote an opposition article to the U.S. occupation of the Philippines?
   - William Jennings Bryan
   - Albert Beveridge
   - Calvin Coolidge
   - William Taft

2. True or False: The Filipino people were working toward their own self-government and independence.
   - True
   - False

3. Which of the following was an argument in favor of invading the Philippines from the reading?
   - Education of the Filipino people
   - Strategic military geography
   - Fruit importation
   - To get oil

4. The political cartoon titled, "Showing Light to the Filipino People," represents which of the following arguments?
   - The American people planned to introduce electricity to the Philippines.
   - American control of the Philippines would help America grow economically.
   - Americans would help the Philippines become more "civilized?"
   - America and the Philippines would work together towards prosperity.

5. True or False: The American government tried to portray the Philippines as incapable of self-government in order to rationalize invading the Philippines, even though it was not an accurate depiction.
   - True
   - False
6. Both texts mention foundational American principles and documents (i.e. The Declaration of Independence). How do each of the authors of the texts use these principles and documents in their argument?

- Both authors use these founding principles to argue in favor of invading the Philippines.
- Both authors use these founding principles to argue in opposition to invading the Philippines.
- Bryan argued these principles do not prohibit invading, while Beveridge argued that invading goes against our founding principles.
- Beveridge text argued these principles do not prohibit invading, while Bryan argued that invading goes against our founding principles.

7. Which of the following is the best caption for the photograph of the soldiers?

- "American soldiers invade Filipino streets."
- "Joint forces march to celebrate American occupation."
- "Filipino troops practice military drills in streets."
- "Firefight erupts in the Philippines."

8. Which of the following statements best summarizes the exact reason for American occupation of the Philippines?

- It was American duty to spread democracy across the world in order to lift up other countries in need such as the Philippines.
- Americans only wanted to occupy the Philippines in the interest of strategic military geography, raw goods, and land.
- The Philippines wanted the assistance of the United States to help them establish a government and economy.
- The United States wanted to spread its influence to become a world power under the guise of assisting the Filipino people.
9. How do the two political cartoons relate to each other?

- They share the same stance in support of invading the Philippines.
- One shows how the U.S. will help the Philippines, while the other shows the cruelty of imperialism.
- They are completely unrelated to each other.
- They share the same stance of how imperialism will hurt the Philippines.

10. Which of the following statements would both authors of the text agree with?

- God wills the United States to occupy the Philippines as Manifest Destiny.
- It is wrong to do unto other countries what England had once did unto the American colonies.
- The Filipino people are not educated.
- Filipino history is rich in culture and arts.
Appendix B - Experimental Documents

Document 2: Albert Beveridge Defends U.S. Imperialism, 1900
Document 3: Photograph 1 American Soldiers in the Philippines
Document 4: Political Cartoon 1 United States Imperialism
Document 5: Political Cartoon 2 Showing Light to the Filipino People
Document 6: Photograph 2 Filipino Nationalists Work Toward Independence
Document 1: William Jennings Bryan Opposes U.S. Occupation of the Philippines, 1900

The young man upon reaching his majority can do what he pleases. He can disregard the teachings of his parents; he can trample upon all that he has been taught to consider sacred; he can disobey laws of the State, the laws of society and the laws of God. He can stamp failure upon his life and make his very existence a curse to his fellow men, and he can bring his father and mother in sorrow to the grave; but he cannot annul the sentence, "The wages of sin is death."

And so with the nation. It is of age and it can do what it pleases; it can spurn the traditions of the past; it can repudiate the principles upon which the nation rests; it can employ force instead of reason; it can substitute might for reason; it can substitute might for right; it can conquer weaker people; it can exploit their lands, appropriate their property and kill their people; but it cannot repeal the moral law or escape the punishment decreed for the violation of human rights...

Some argue that American rule of the Philippine Islands will result in the better education of Filipinos. Be not deceived. If we expect to maintain a colonial policy, we shall not find it to our advantage to educate the people. The educated Filipinos are now in revolt against us, and the most ignorant ones have made the least resistance to our domination. If we are to govern them without their consent and give them no voice in determining the taxes they must pay, we dare not declare them, lest they learn to read the Declaration of Independence and Constitution of the United States and mock us for our inconsistency.

Document 2: Albert Beveridge Defends U.S. Imperialism, 1900

MR. PRESIDENT, the times call for candor. The Philippines are ours forever, "territory belonging to the United States," as the Constitution calls them. And just beyond the Philippines are China's illimitable markets. We will not retreat from either. We will not repudiate our duty in the archipelago. We will not abandon our opportunity in the Orient. We will not renounce our part in the mission of our race, trustee, under God, of the civilization of the world. And we will move forward to our work, not howling out regrets like slaves whipped to their burdens but with gratitude for a task worthy of our strength and thanksgiving to Almighty God that He has marked us as His chosen people, henceforth
to lead in the regeneration of the world.

This island empire is the last land left in all the oceans. If it should prove a mistake to abandon it, the blunder once made would be irretrievable. If it proves a mistake to hold it, the error can be corrected when we will. Every other progressive nation stands ready to relieve us.

But to hold it will be no mistake. Our largest trade henceforth must be with Asia. The Pacific is our ocean. More and more Europe will manufacture the most it needs, secure from its colonies the most it consumes. Where shall we turn for consumers of our surplus? Geography answers the question. China is our natural customer. She is nearer to us than to England, Germany, or Russia, the commercial powers of the present and the future. They have moved nearer to China by securing permanent bases on her borders. The Philippines give us a base at the door of all the East.

Lines of navigation from our ports to the Orient and Australia, from the Isthmian Canal to Asia, from all Oriental ports to Australia converge at and separate from the Philippines. They are a self-supporting, dividend-paying fleet, permanently anchored at a spot selected by the strategy of Providence, commanding the Pacific. And the Pacific is the ocean of the commerce of the future. Most future wars will be conflicts for commerce. The power that rules the Pacific, therefore, is the power that rules the world. And, with the Philippines, that power is and will forever be the American Republic.

But if they did not command China, India, the Orient, the whole Pacific for purposes of offense, defense, and trade, the Philippines are so valuable in themselves that we should hold them. I have cruised more than 2,000 miles through the archipelago, every moment a surprise at its loveliness and wealth. I have ridden hundreds of miles on the islands, every foot of the way a revelation of vegetable and mineral riches.

Here, then, senators, is the situation. Two years ago there was no land in all the world which we could occupy for any purpose. Our commerce was daily turning toward the Orient, and geography and trade developments made necessary our commercial empire over the Pacific. And in that ocean we had no commercial, naval, or military base. Today, we have one of the three great ocean possessions of the globe, located at the most commanding commercial, naval, and military points in the Eastern seas, within hail of
India, shoulder to shoulder with China, richer in its own resources than any equal body of land on the entire globe, and peopled by a race which civilization demands shall be improved. Shall we abandon it? That man little knows the common people of the Republic, little understands the instincts of our race, who thinks we will not hold it fast and hold it forever, administering just government by simplest methods...

But, Senators, it would be better to abandon this combined garden and Gibraltar of the Pacific, and count our blood and treasure already spent a profitable loss, than to apply any academic arrangement of self-government to these children. They are not capable of self-government. How could they be? They are not a self-governing race. They are Orientals, Malays, instructed by the Spaniards in the latter's worst estate...

The Declaration of Independence does not forbid us to do our part in the regeneration of the world. If it did, the Declaration would be wrong, just as the Articles of the Confederation, drafted by the very same men who signed the Declaration, was found to be wrong. The Declaration has no application to the present situation. It was written by self-governing men for self-governing men...

...This is our divine mission of America, and it holds for us all the profit, all the glory, all the happiness possible to a man. We are trustees of the world's progress, guardians of its righteous peace. The Judgement of the Master is upon us: "Ye have been faithful over a few things; I will make you ruler of many things."

Document 3: Photograph 1 American Soldiers in the Philippines

Document 4: Political Cartoon 1 United States Imperialism
Document 5: Political Cartoon 2 Showing Light to the Filipino People

![Political Cartoon 2](image)

Document 6: Photograph 2 Filipino Nationalists Work Toward Independence

![Photograph 2](image)
Appendix C - MDQI Instruction

Introduction
Video Transcript
Example Question Frameworks
Introduction

Thinking critically about information often requires making sense of information that comes from different sources and those sources don't always agree. A good question you ask can help you to start the critical thinking process of synthesizing, analyzing, and evaluating those multiple documents. History materials provide a good opportunity to learn how to do this because of the multiple documents that can be taken from different historical perspectives. In this section you will learn how to develop good questions in history that will help you to synthesize, analyze, and evaluate the multiple documents you will read in the following section.

Video Transcript

Two students decide to go skiing for the weekend and are having such a good time they decide to blow off their calculus exam in order to get some final runs in. They decided that they will tell their professor that they missed the exam because they got a flat tire, and deserve to retake the exam. Hearing the story, the professor agrees that it was just bad luck and of course they can take the exam later. At the designated time, the professor greets the students and puts them into two separate rooms to take the exam. The first few questions are worth 10% of the exam and are pretty easy. The students grow progressively more confident, sure that that they have gotten away with fooling the professor into letting them take this exam. However, when they turn to the second page they discover that they really haven’t. The last question is worth 90% of their grade and reads, “Which tire?”

What this professor was trying to do was to check credibility of the story and to better understand what happened by asking a question. We need to do the same thing when studying
multiple documents. A single source of information is not enough to tell the whole story, so like the professor, we must ask questions to understand the whole story.

When studying multiple documents, you might ask yourself, “Who wrote this?” or “Why did they write this?” or “What is the main idea?” All of these questions help you to understand the author’s point of view or the main idea of a single text. It does not help you to understand this document in comparison to other texts to see inconsistencies or similarities across the documents.

Again the purpose of studying multiple documents is that we may not be able to get the full story from a single document alone, so we must analyze and synthesize multiple documents to get the full story. We must then ask questions that help us to analyze and synthesize multiple documents, and not just stop at the single document questions.

Think about writing an essay for class in which you would be graded on the number of ideas and concepts you included in the response. Would you rather have, “What is an unconditioned stimulus according to Pavlov?” or “Compare and contrast the major learning theories and their contributions to understanding teaching and learning?”

The second question allows you to connect a greater number of ideas from multiple documents. The first simply asks you to recall information from a single source like your textbook, without making any connections to any other important ideas from class.

This concludes the instructional video for learning how to generate multiple document questions. You will now go on to practice identifying multiple document questions before going on to generate your own multiple document questions while studying a multiple document set on school choice.
Example Question Frameworks

Below are some generic question frameworks to help guide your understanding of higher-level questions in history. Remember, the purpose is to find patterns and differences in and between documents. When you read through these question examples, consider how you could use your Historical Document Organization Matrix to generate these types of questions.

How do the points of view differ in… and… and what does this mean for (time period)?

How does… affect…?

How does… differ from the traditional points of view from history?

What are the consistencies and/or the inconsistencies in the documents?

Are these accounts consistent with other accounts of the time period?

From what lenses can we view this historical event given these accounts?
Vita
Chelsea E. Cameron

Education

M.S. Educational Psychology, 2015, The Pennsylvania State University, University Park, PA
Thesis: Fostering Historical Thinking and Integration: The Effects of Instruction on Student Question Generation in History

B.S. Secondary Education, Social Studies, 2011, The Pennsylvania State University, University Park, PA
With Honors and Distinction, Honors Thesis: A Taxonomy of Instructional Representations in Social Studies Instruction

Professional Experience

Adjunct Faculty, College of Education, Duquesne University, Pittsburgh, PA

Academic Support for Student Athletes Intern, Duquesne University, Pittsburgh, PA

Course Instructor, The Pennsylvania State University, University Park, PA

Research Assistant, NSF Grant, Quality Talk Project, Department of Educational Psychology, University Park, PA

Research Assistant, HHMI Grant: Biology Curriculum Assessment, Department of Biology, The Pennsylvania State University

Teacher Assistant, The Craig School, Mt. Lakes, NJ

Publications

Van Meter, P.N. & C. Cameron (In Press). The effects of representation on multiple document notetaking. Learning and Instruction

http://doi.org/10.1016/j.learninstruc.2017.01.003


Selected Presentations


Sperling, R.; Cameron, C.; Follmer, D.J.; Reeves, P. (August 2017) The interaction of task and levels of prior knowledge on the comprehension of multiple documents. Paper presented at the annual meeting of the American Psychological Association, Washington, D.C.
