International Electronic Journal of Elementary Education, 2012, 5(1), 63-78.

The Role of CLEAR Thinking in Learning Science from Multiple-Document Inquiry Tasks

Thomas D. GRIFFIN*
University of Illinois at Chicago, Chicago, IL, USA

Jennifer WILEY
University of Illinois at Chicago, Chicago, IL, USA

M. Anne BRITT
Northern Illinois University, Dekalb, IL, USA

Carlos R. SALAS
University of Illinois at Chicago, Chicago, IL, USA

Received: 21 August 2012 / Revised: 6 November 2012 / Accepted: 6 November 2012

Abstract

The main goal for the current study was to investigate whether individual differences in domain-general thinking dispositions might affect learning from multiple-document inquiry tasks in science. Middle school students were given a set of documents and were tasked with understanding how and why recent patterns in global temperature might be different from what has been observed in the past from those documents. Understanding was assessed with two measures: an essay task and an inference verification task. Domain-general thinking dispositions were assessed with a Commitment to Logic, Evidence, and Reasoning (CLEAR) thinking scale. The measures of understanding were uniquely predicted by both reading skills and CLEAR thinking scores, and these effects were not attributable to prior knowledge or interest. The results suggest independent roles for thinking dispositions and reading ability when students read to learn from multiple-document inquiry tasks in science.

Keywords: Thinking Dispositions, Learning From Text, Climate Change, Earth Science, Multiple-Document Inquiry Tasks.

Introduction

The internet has become a primary means by which people search for information to answer many science-related questions. Adults read internet sources to help them understand phenomena in the world around them. They read to learn about the development of new

* Thomas D. Griffin, Department of Psychology, University of Illinois at Chicago, 1007 W. Harrison St (M/C 285), Chicago, IL, 60607 USA, Phone: (312) 996-8190, E-mail: tgriffin@uic.edu.

ISSN:1307-9298

www.iejee.com

technologies. They read to form beliefs on policy issues as well as to inform health-related decision-making. To form an understanding, they need to integrate information that is spread out across numerous documents and sources. Hence, the ability to learn about scientific phenomena from multiple documents is a critical skill for life-long learning. Yet, instruction in this area is not well represented in science classrooms, and research identifying potential sources of variance in how well students can engage in multiple-document inquiry learning on science topics has been limited. The main goal for the current study was to investigate whether individual differences in either reading skills or domain-general thinking dispositions might affect learning from multiple-document inquiry tasks in science.

What Processes are needed to Learn from Multiple-Document Inquiry Tasks?

One reason why learning from multiple documents is so complex is because it requires all of the processes necessary for comprehending individual informational texts, plus an additional set of processes that become particularly important when readers are confronted with information from more than one text. According to theories of text comprehension (Kintsch, 1998; Kintsch & van Dijk, 1983), understanding even a single informational text requires the construction of several levels of representation. At the lowest level, a reader creates a surface representation, which generally consists of a fleeting episodic trace capturing the exact words and format of the text. At the next level of processing, the reader attempts to develop the text-based representation. This is essentially a propositional representation of the ideas presented in each clause or sentence. Basic word and sentence-level reading processes contribute to the construction of this text-based representation. In addition, to learn from informational text, the reader must attempt to develop yet another level of representation, referred to as the situation model by Kintsch. On this level, the reader attempts to connect ideas between the sentences and with prior knowledge to develop a coherent understanding of the content that is being described. When the goal for reading informational science texts is to develop an understanding of how or why a phenomenon occurs, then the situation model can be thought of as a causal chain or mental model of the phenomena being described (Kintsch, 1994; Trabasso & van den Broek, 1985; Wiley, Griffin & Thiede, 2005).

Yet, in many learning situations, readers are presented with more than a single text from which to obtain information (Britt, Perfetti, Sandak, & Rouet, 1999; Perfetti, Rouet, & Britt, 1999). One framework for describing the cognitive processes involved in multiple-document comprehension is the MD-TRACE model (Multiple Documents - Task-based Relevance Assessment and Content Extraction) proposed by Rouet and Britt (2011). According to the MD-TRACE model, students begin multiple-document reading by creating an interpretation of the task (called the Task Model). This Task Model includes the goals and subgoals for reading (e.g., why is the text being read? what is the question to be answered? what does developing an argument or explanation entail?) and plans to reach those goals (e.g., find evidence or causes). In other words, the task model includes the goals for reading and the basic steps that should be taken to achieve the desired outcome. Depending on the reading context, including the reader's interpretation and the instructions that are given, task models will range from cursory to clearly delineated. Although reading goals may affect learning in single text contexts, they become even more critical to consider in multiple-document contexts.

Learning from multiple documents instead of a single document also requires another level of representation (a Documents Model) that captures the relation of information across the document set, and information about each document, in addition to the representations of the content of individual texts. The Documents Model (Britt, Perfetti, Sandak, & Rouet, 1999;

Perfetti, Rouet, & Britt, 1999) has been proposed to capture these two needs. One part of a Documents Model, the Intertext Model, includes information about the sources of the various documents (e.g., who wrote it), and notes relations among the documents (e.g. the presence of corroborating or conflicting statements). The Documents Model also contains the Integrated Model, which serves as the representation of the situation or phenomena described across documents. According to the MD-TRACE model, the extent to which readers develop Integrated Models or Intertext Models from multiple-document contexts will be partially determined by the goals readers have in their Task Model.

With a single text, comprehension can be driven by a text's intended purpose, structure, or argument. With multiple texts, the reader must impose selection and organization in order to form a model that integrates the information from different texts (rather than simply constructing distinct models of each text). A reader's goals guide the process of reading and evaluating the individual texts, selecting relevant information, and reassembling what is selected into a new coherent model. Thus, the interpretation of the task and the goals a reader sets for reading are a critical determinant of multiple-document comprehension.

What Leads to Successful Learning from Multiple-Document Inquiry Tasks in Science?

A burgeoning area of investigation at the intersection of literatures on subject-matter learning and learning from informational texts has been exploring what conditions facilitate student understanding from multiple-document inquiry activities in science (Braten, Britt, Stromso,& Rouet, 2011; Cerdan & Vidal-Abarca, 2008; Goldman, Braasch, Wiley, Graesser, & Brodowinska, in press; Mason, Boldrin, & Ariasi, 2010; Sanchez, Wiley & Goldman, 2006; Wiley, Ash, Sanchez & Jaeger, 2011; Wiley, Goldman, Graesser, Sanchez, Ash, & Hemmerich, 2009). There are numerous factors that are likely to impact learning from multiple documents. These factors can include features of the set of sources that are provided, as well as the nature of the inquiry task that is given. One general approach within this literature has been to provide students with a set of informational texts as reading material, often through the guise of the results of an internet search. As a goal for the processing of the informational texts, students are generally tasked with learning about how or why a phenomenon occurred such as "What caused the eruption of Mt. St. Helens?" or "How do bacteria resist the effects of antibiotics and which biological mechanisms explain this phenomenon and its transmission to other bacteria?" or "What caused the extinction of dinosaurs?" Similar to studies on learning from multiple sources in history (i.e. Wiley & Voss, 1996; 1999), when students are prompted to use the text sets to generate a causal argument or explanatory model of a phenomenon, it results in better learning from the activity. For example, Cerdán and Vidal-Abarca (2008) found that prompting students to read texts in order to explain how resistance to antibiotics develops resulted in a deeper and more integrated understanding than did asking students very specific questions that could be answered by searching for, finding, memorizing, and reproducing isolated bits of information within the text set. These results are generally consistent with the idea that students' understanding may benefit from multiple-document activities to the extent that students engage in constructive processing that builds connections across ideas in order to form a coherent, integrated model of the phenomena (Britt & Rouet, 2012; Wiley & Voss, 1996; 1999).

Yet, further research has demonstrated that not all learners take advantage of this opportunity, especially when the reading material requires selective use of information (Wiley, Ash, Sanchez & Jaeger, 2011). For example, in Wiley et al (2009) which provided students with texts from both reliable and unreliable sources about volcanic eruptions, the ability to evaluate the sources was seen as a gatekeeper to the development of an accurate

mental model of the phenomena. In the first study, better learning was related to the ability to differentiate reliable from unreliable sources. In the second study, the presence of a pre-inquiry instructional module on source evaluation was manipulated (see also Sanchez, Wiley, & Goldman, 2006). In this module, called "SEEK", students were taught to evaluate reliability of each document by considering not only the source of each document, but also whether any evidence was presented that could be related to an explanation of the phenomenon prompted by the inquiry question, and also how information in the document related to other knowledge about the phenomena. Participants completed the SEEK module on an unrelated topic prior to the inquiry task. Thus, this instructional manipulation stressed the need to consider the reliability of Source information as well as the importance of thinking about Explanations, Evidence, and integration with both prior Knowledge and the information in other documents. The main result of this study was that students given SEEK instruction demonstrated better learning from the subsequent multiple-document inquiry task on volcanic eruptions. Together these two studies demonstrated that multiple-documents learning is predicted by individual differences in evaluating sources and content of the documents, and that learning is improved by instruction that targets skills related to evaluation based on source, evidence, and coherence with other information.

Additional lines of investigation have further explored what the better learners were doing in the first Wiley et al (2009) study by using eyetracking and think-aloud methodologies. Wiley, et al. (2011) found that when students were asked to write an argument explaining the causes of volcanic eruptions, the best learners showed more selective reading behaviors. These participants were more likely to skim pages, but would often go back and thoroughly re-read a page if it contained conceptually relevant information. Eyetracking data showed that these better learners were also more likely to spend a greater proportion of their time on the specific sentences of a relevant page that (i.e., regions of interest, ROIs) were most critical for forming an explanation. Lastly, while all readers looked at the illustrations that accompanied the texts, the better learners tended to look at the conceptual images more so than decorative images. These findings suggest that the students that were more goal-directed, strategic, and selective in their reading and use of available information learned more and created better causal explanations. These better learners seemed to be more engaged in the process of creating an integrated mental model, as they showed better use of conceptual illustrations, and more integrated reading patterns. Importantly, the best learners were not simply spending more overall time reading—but instead they were more selective with their reading efforts. For the best learners, it appears that they responded to an argument-writing task by directing their attention to the most relevant information for the construction of an accurate mental model or explanation. An additional finding in this study is that when students were given an instruction to write a report rather than an argument, they were generally less selective in which information they read and included in their essay.

The findings of the think-aloud study (Goldman, et al. in press) where readers were asked to simply think aloud about what they were doing while viewing the documents, also suggested that better learners were more selective in what they read and how they utilized the information. Better learners made more comments related to their evaluation of source credibility and reliability, especially in relation to why they went to certain pages and not others or why they were leaving a page before they read all of its content. Their comments revealed a more strategic approach to reading in which they referenced their inquiry task, their current understanding, and what they still needed to accomplish. Consequently, they incompletely read pages that they judged would not further their understanding and finished pages that they judged would further their understanding. In addition, the better

learners were more selective in how they used the information. They engaged in more sensemaking, self-explanation and comprehension-monitoring processes while they were reading reliable sites than they did on unreliable websites. These behaviors are usually related to incorporating information into mental models. Self-explanation helps in both the construction of models (Chi, 2000) and as a source of cues for evaluating the quality and completeness of one's own mental model (Griffin, Wiley, & Thiede, 2008). Thus, these findings suggest that better readers were more selective and integrated the more reliable information into their mental models.

Interestingly, the think-aloud evidence available from this study does not suggest that the better learners had better a priori knowledge of which sites might be considered more reliable and useful. Rather, the think-aloud comments suggested that readers assessed whether their understanding was increasing and what additional information was needed to achieve the inquiry goal. This type of evaluation of the sites in terms of goal-relevant content resulted in learners being more strategic in their reading and spending a greater proportion of their time on the more reliable sites, hence resulting in better learning from the inquiry task.

Other work on learning from internet-inquiry tasks also suggests that differences in how students approach the evaluation of information quality can affect their learning. Following an internet inquiry task on dinosaur extinction, Mason, Boldin, and Ariasi (2010) asked learners how they decided which information they found on the internet was true. They found that students who were more likely to appeal to comparisons of information from multiple sources and to scientific evidence as a basis for evaluation were those who were more likely to learn the scientifically accepted conception of extinction from the activity. This study also administered the Conley, Pintrich, Vekiri and Harrison (2004) self-report measure of epistemic beliefs about science. They found that scores on the Justification subscale were related to whether students learned the scientifically accepted view of extinction. This subscale assesses beliefs about the nature of science and the importance science places on experiments, replication, and the source of scientific ideas. Similarly, other work from Braten and Stromso (2010) has also shown that some features of epistemological beliefs (specifically about the to-be-learned topic such as climate change) can predict who might learn most effectively from multiple-document inquiry tasks, and that readers who engage in more source evaluation behaviors develop better understanding from multiple documents (Braten, Stromso & Britt, 2009).

Together, these results suggest that multiple-document inquiry tasks provide the opportunity for readers to engage in more integrative processing and model construction. Yet, the extent to which readers are able to take advantage of this opportunity depends on whether readers selectively process and integrate the most reliable, central, and relevant information. Several lines of research suggest that individuals differ in how they approach inquiry tasks and whether they engage in evaluation, selective reading and integration. The reviewed research suggests that beliefs about science matter, as do instructions to use the information in order to form causal arguments and instructions to evaluate evidence, relevance, and source reliability. Individual differences in general thinking dispositions may be a promising source of influence on the likelihood that a person will engage in successful learning from multiple documents.

Delineation of two classes of individual differences: Capacities and Dispositions.

Expanding on the work of reasoning theorists such as Baron (1985), Stanovich and West (1997) delineated two classes of individual differences, capacities and dispositions that may be useful to consider in this context. What and how a person does on any cognitive task is determined by a combination of what they are capable of doing and what they are disposed towards doing. Capacity constraints have been a primary focus within Cognitive Psychology. These range from more basic processing abilities and constraints that are not likely teachable, such as working memory capacity, to more teachable factors including reading skills such as decoding and word knowledge. Individual differences in capacity constraints would be expected to play a role in learning from multiple documents just as they would in learning from single passages. Although rarely included in studies on learning from multiple documents, reading ability has been shown to be a significant predictor of multiple-document comprehension in at least one study (Mason, Boldrin, & Ariasi, 2010).

In contrast, dispositional individual differences are those that relate to a person's goals, their orientation towards the task, and their willingness towards applying whatever relevant skills and capacities they have to the processes required for effective thinking, learning, and task performance. Griffin and Ohlsson (2001) showed that people vary in whether they report forming their belief on a topic in terms of considering relevant evidence versus deferring to their affective preferences, and this in turn predicts people's willingness to revise their belief in light of new evidence. More recently, Griffin (2008) reported findings which suggest that people have a general disposition towards whether they consider evidence or affect when forming their beliefs. People reported the extent to which they based their beliefs on either considering evidence or relying upon faith. The beliefs varied across eight different topics that were both religious and non-religious. The topics were largely unrelated in content as indicated by the fact that *what* a person believed on each topic did not predict what they believed on the other topics. However, the degree to which people relied upon faith versus evidence to arrive at whatever belief they held on a topic was correlated with their reliance on faith versus evidence for all other topics. This pattern of consistency across distinct topics implicates a general thinking disposition relating to intellectual values.

A reader's goals when attempting to learn from multiple documents might be affected by many factors, including their interest in the topic, as well as by general intellectual values. In a multiple-documents inquiry context, effective learning requires engaging in evidence-based reasoning in the service of argument construction. A general disposition of valuing evidence-based thinking in the evaluation of beliefs and claims would seem to orient one toward attempting to construct the kind of coherent argument that such an inquiry task requires. Thus, having an evidence-based disposition may play a unique role above and beyond learners' capacities and skills in determining whether they engage in the extra-textual processing required for developing an integrated model of phenomena across texts as required by multiple-document inquiry tasks.

In previous studies on domain-general thinking dispositions, Stanovich and colleagues have shown that a general disposition towards actively open-minded thinking (AOT) predicts cognitive performance on higher-order thinking tasks over and above measures of cognitive capacity, such as the SAT, Raven's Matrices, and Nelson-Denny Reading Comprehension Test (e.g., Stanovich & West, 1998; for a recent review see Stanovich, 2012). This research has employed various versions of a 41-item AOT scale comprised of several subscales. Some items tap moral authoritarianism and openness to others' values (e.g., "I believe we should look to our religious authorities for decisions on moral issues."; "There are a number of people I have come to hate because of the things they stand for."). However, other items

focus upon a general openness to intellectual inquiry, evidence, and belief revision (e.g., "People should always take into consideration evidence that goes against their beliefs."; "One should disregard evidence that conflicts with your established beliefs."). The present research is interested in these latter items assessing a disposition that is more directly related to the kind of evidence-based thinking that should impact a multiple-documents inquiry task in science. The AOT scale has typically been employed as a composite with a single score that is used to predict performance on higher-order cognitive tasks assessing logical reasoning, rational judgment, normative decision making, and informal reasoning processes such as syllogistic reasoning, probabilistic reasoning, statistical reasoning, covariation detection, and argument evaluation (Stanovich & West, 1997; 1998). Individuals with high AOT scores evaluate objective argument quality more accurately than those with lower AOT scores, and their evaluations are less biased by consistency with prior beliefs, even when controlling for cognitive ability (Stanovich & West, 1997; West, Stanovich, & Toplak, 2008). Sá, Kelley, Ho, and Stanovich (2005) found that people low in AOT were more likely to generate arguments that simply reiterated their personal theory rather than providing supporting evidence. Identifying and incorporating relevant information across multiple texts is likely to be impacted by some of the same factors that impact the kind of argument evaluation and construction tasks employed in these studies.

Although most of the studies on AOT have involved adults, one study has found that thinking dispositions can predict performance on several standard reasoning tasks with children (Kokis et al., 2002). However, none of the prior studies have used the AOT measure to examine the effects of thinking dispositions in a classroom learning task where the goal is to acquire knowledge and understanding in a content area. Prior research shows that this thinking disposition impacts performance on tests of one's reasoning proficiency, but has not examined the impact on the content learning and knowledge acquisition that partially depend on such skills. Thus, it is informative to examine individual differences in this thinking disposition in a real classroom context where the goal of the task is science learning via selective integration of information across multiple information sources and inclusion of that information in the form of an explanatory argument.

The Present Study

The purpose of the current study was to explore the effects of individual differences in both capacity and dispositional constraints on learning from a multiple-document inquiry task in science. Middle school students were given a set of documents about the global temperature system and were asked to write an essay explaining how and why recent patterns in global temperature are different from what has been observed in the past. Understanding of the science topic was assessed both by considering the quality of the essays that were written as well as by performance on an inference verification task. The main question for the current study was whether individual differences in domain-general thinking dispositions might have unique effects from reading skill on the understanding that results from a multiple-document inquiry task in science.

Method

Participants

Participants in this study were 59 seventh grade students from 3 science classes in an urban public middle school in the United States. The average age was 13.31 years (*SD* = .64). The sample was 57% female. Self-reported ethnicity was 22% Hispanic, 27% African American,

10% Asian, 59% White, 30% Native American/Pacific Islander and 25% Other. (Students were able to select multiple ethnicities and 5 students did not select any.)

Materials and Measures

Global Temperature Document Set. All participants were given a set of 7 documents containing information related to the causes of global temperature change, based on material that has been used in previous studies with older students about the causes of Ice Ages (Sanchez & Wiley, 2006; Sanchez & Wiley, 2009). Five text-based documents covered several main topics including Ice Ages, the Carbon Cycle, The Greenhouse Effect, Solar Radiation, and Energy from Fossil Fuels. The document set also included a graph of CO^2 Concentrations over the last 400,000 years, presented as its own document. In addition, students were provided with seventh document, titled "Changes in Global Temperatures", which provided textual background on the methods used to assess global temperatures. This document also included a graph of average global temperatures over the last 400,000 years, and a second graph showing the increases in average global temperatures from 1870 to 2010.

The texts were excerpted from several online sources from the United States Geological Survey, the Public Broadcasting Service, the NASA earth observatory, the Environmental Protection Agency, as well as an extension module from an earth science textbook series (Bennington, 2009). To adapt the texts for younger grade levels, vocabulary and sentence structures were simplified. The final text-based documents were on average 326 words long (range: 208-475), with an average in Flesch Reading Ease of 62.36, and an average Flesch-Kincaid grade level of 7.9. The documents were presented to students on pieces of paper contained in a pocket folder, with each of the 7 documents printed on a separate page.

Inquiry Task Essay Prompt and Essay Coding. One main source of information about student understanding was the essays that students wrote in response to the inquiry prompt. Students were asked to "use this set of documents to write an essay explaining how and why recent patterns in global temperature are different from what has been observed in the past." Student responses to this essay prompt were evaluated for the presence of 5 critical target concepts that directly relate to recent changes in global temperature, and thus address the inquiry question students were asked. These concepts were:

1. We are in an unusually long warming period.
2. CO^2 levels in the atmosphere are at their highest in at least 400,000 years.
3. Fossil fuel burning releases CO^2.
4. CO^2 is a greenhouse gas.
5. Greenhouse gases in the atmosphere cause warming.

All essays were evaluated for the presence of the target concepts by two independent coders, who produced a high level of interrater reliability (Krippendorf's $\alpha = .90$, $p < .05$). Any differences were resolved through discussion.

Inference Verification Task. As another measure of student understanding, a sentence judgment task was created in which students were asked to indicate which of a list of statements seemed true based on the texts they had just read. This test (based on Sanchez & Wiley, 2006; Wiley & Voss, 1999) consisted of 18 statements that represented potential connections or inferences that could or could not be made based on the information in the document set. Some example items are "In the past 100 years, both fossil fuel use and CO^2 levels have increased" and "Increases in fossil fuel use increase the amount of heat that escapes into space." The first is an example of a conclusion that is supported by the documents but requires connections across documents. The second is an example of a

statement that is false based upon connecting multiple ideas across documents, namely ideas 3, 4, and 5 that were coded in the essays. The items represented 8 correct and 8 incorrect inferences. For every correct and incorrect inference appropriately identified, the students received a single point. An overall proportion score was computed for the task, and higher levels of performance indicated better understanding of the inferences that could be made from the documents.

CLEAR Thinking Scale. Students' CLEAR thinking refers to their Commitment to Logic, Evidence, and Reasoning. The 5-item scale assesses the extent to which students place value and importance on reasoning about evidence when forming and revising beliefs. The construct is measured at the most domain general level. The scale incorporates items from the flexible thinking scale (Stanovich & West, 1997) and the belief identification scale (Sa', West, & Stanovich, 1999), which were revised by Kokis et. al.,(2002) to be used with children. The items were selected based upon the criteria that they directly ask about belief revision in the face of new evidence or information.

The items used for the CLEAR Thinking Scale were:

1. I never change what I believe in - even when someone shows me that my beliefs are wrong.
2. People should always consider evidence that goes against their beliefs.
3. It's important to change what you believe after you learn new information.
4. People shouldn't pay attention to evidence that contradicts their strongly held beliefs.
5. To decide what is true, you often have to ignore your emotions and stick just to the evidence.

Students were asked to respond to these items using a 1-6 scale with 1 meaning Strongly Disagree and 6 meaning Strongly Agree. The scale was scored by subtracting the average rating for the negatively worded items (1 and 4) from the average of the positively worded items. A difference score between the weighted averages means that the combined influence of the positively worded items on the total score is the same as the combined influence of the negatively worded items. This avoids the problem of giving more weight to positively worded items as a group, which creates a response bias effect such that people who simply anchored all of their ratings at a higher value would receive a higher score.

Descriptive Student and Teacher Surveys. Due to the Family Educational Rights and Privacy Act (FERPA), official student standardized test scores could not be obtained. To obtain measures of reading skill in lieu of test scores, a teacher survey asked teachers to indicate each student's level of reading skill relative to their grade level as low, medium or high. A student self-report survey was created to collect basic descriptive information including gender, date of birth, and ethnicity. Students were also asked to rate on a 1-to-5 scale their level of interest in science, interest in the topic, and prior knowledge about the topic.

Procedure

Students participated in the inquiry activity as part of their normal science classes. All materials for the inquiry activity were distributed to students in folders, including the inquiry task essay prompt, blank writing pages and the document set. Students were asked to read along as the inquiry task essay prompt instruction was read out loud.

The full instructions for the Reading and Writing task were:

The primary purpose of reading in science is to understand the causes of scientific phenomena. This means your goal for reading is to understand how and why things happen. To reach an understanding of a new topic in everyday life, we often need to gather information from multiple sources. In today's task your goal is to learn about the causes of global temperature changes from several documents. You will have to piece together important information across the documents to construct a good understanding. No one text will provide the answer. This task is interesting because you are the one making the connections across documents and coming up with an explanation. No author has already done the work for you. It is also important that you use information from the documents to support your explanation of the causes.

Your task is to use this set of documents to write an essay explaining how and why recent patterns in global temperature are different from what has been observed in the past. Be sure to use specific information from the documents to support your conclusions and ideas.

Students had access to the documents as they wrote the essays. Then, the essays and document sets were collected and students completed the Inference Verification Task, without access to the documents. These were collected and students completed a final booklet including the CLEAR Thinking Scale and the self-report descriptive student surveys. Teachers were asked to fill out the teacher survey while students worked on the inquiry task. The activity was done over two 50 minute periods.

Results

Descriptive Statistics

The descriptive statistics for all variables in the study are displayed in Table 1. All variables showed normal distributions and high variance covering the range of possible values. The mean CLEAR Thinking score was greater than 0, reflecting that most students had at least slight agreement with an evidence-based disposition. However, there was high variability and many students had negative scores and disagreed with an evidence-based disposition.

Table 1. *Means, Standard Deviations, and Ranges for all Measures*

Measure	Mean	SD	Observed Range	Possible range
CLEAR Thinking	1.19	1.93	-2.67 - 4.67	-5.00 - 5.00
Reading Skill	2.32	0.78	1.00 - 3.00	1.00 - 3.00
Prior Knowledge	3.46	1.18	1.00 - 5.00	1.00 - 5.00
Science Interest	2.98	1.42	1.00 - 5.00	1.00 - 5.00
Topic Interest	2.88	1.25	1.00 - 5.00	1.00 - 5.00
Essay Concepts	1.83	1.66	0.00 - 5.00	0.00 - 5.00
Inference Test	0.70	0.14	0.44 - 1.00	0.00 - 1.00

Correlations among Measures of Understanding and Individual Differences

As shown in Table 2, there was a significant positive correlation between the two outcome measures of understanding. Students with greater conceptual coverage in their essays (Essay Concepts) also tended to have higher scores on the Inference Verification Task (Inference Test), despite the fact that the texts were only available during the essay writing. Table 2 also shows that both these measures of understanding were predicted by CLEAR Thinking dispositions and by reading skill, and that prior knowledge predicted inference test performance. The relationship between prior knowledge and essay concepts was trending in the same direction but weaker and non-significant ($p = .15$).

Table 2. *Pearson Correlations among CLEAR Thinking, Reading Skill, Prior Knowledge, Science Interest, Topic Interest, Number of Essay Concepts, and Inference Test Scores*

Measure	1	2	3	4	5	6
1. CLEAR Thinking	-					
2. Reading Skill	.26*	-				
3. Prior Knowledge	.27*	.27*	-			
4. Science Interest	.11	.01	.08	-		
5. Topic Interest	.13	.05	.01	.73**	-	
6. Essay Concepts	.36**	.46**	.19	.13	.07	-
7. Inference Test	.39**	.42**	.30*	.13	.14	.45**

Note. $N = 59$. *$p < .05$. **$p < .01$.

Students' interest in both science and the specific topic failed to predict performance on the essays and the inference test, and were also unrelated to CLEAR thinking and the other predictors. However, the two interest measures were highly correlated with each other. In addition, interest levels differed for male and female students. Consistent with prior findings (for a meta-analysis, see Weinburgh, 1995) males had significantly higher interest in both science and the topic (*M*s = 3.32 and 3.42) than females (*M*s = 2.58 and 2.66), *t*s(57) = 2.33 and 2.02, *p*s < .05. Gender did not relate to any of the other predictors or to either outcome measure.

Unique Effects of Thinking Dispositions and Reading Skill on Understanding

The main question for the current study was whether individual differences in domain-general thinking dispositions might have unique effects from reading skill on the understanding that results from a multiple-document inquiry task in science. To examine this question, Reading Skill and CLEAR Thinking scores were entered simultaneously into a regression predicting the number of key explanatory concepts in the essays. As seen in the top half of Table 3, the regression resulted in a significant model accounting for 27% of the variance in Essay Concepts. The beta tests showed that both CLEAR Thinking and Reading Skill each accounted for significant unique variance. The inclusion of key explanatory concepts increased with Reading Skill. In addition, regardless of Reading Skill, students with a stronger general disposition towards evidence-based thinking were more likely to incorporate the key explanatory concepts into their essays.

Table 3. *Regression Analyses Predicting Inference Test Scores and Number of Essay Concepts from CLEAR Thinking and Reading Skill Scores*

Predictor	R^2	*F* Value	B	SEM	*B*	*t* Value
(DV) Essay Concepts						
Model	.27	10.29*				
CLEAR Thinking			.22	.10	.25*	2.13
Reading Skill			.84	.25	.39*	3.31
(DV) Inference Test						
Model	.26	9.81*				
CLEAR Thinking			.02	.01	.30*	2.51
Reading Skill			.06	.02	.34*	2.87

Note. $N = 59$. *$p < .05$.

Another regression was conducted in which Reading Skill and CLEAR Thinking scores were entered simultaneously to predict Inference Test performance. The results reported in bottom half of Table 3 were very similar to the Essay Concepts results. The overall model was

significant and accounted for 26% of the variance in test performance. The beta tests show that both CLEAR thinking and Reading Skill each accounted for significant unique variance. Inference Test performance was better for students with more reading skill. More importantly, regardless of reading skill, students with a stronger general disposition towards evidence-based thinking were more likely to correctly identify statements that could and could not be inferred by integrating the information from the multiple documents. Since prior knowledge of the topic was related to inference performance and CLEAR Thinking (see Table 2), this analysis was rerun adding prior knowledge as a control predictor. The results did not change, except for a slight increase in the total variance explained from 26% to 28%.

Conclusions

Across two measures of student understanding, the results of the present study demonstrate the influence of both reading skill and a domain-general thinking disposition on learning science from multiple-document inquiry tasks. These influences were independent from each other and from self-reported ratings of prior topic knowledge, interest in the topic and interest in science. Of these individual differences, only the interest ratings were not related at all to understanding. Although reading skill and prior topic knowledge were not assessed with standardized measures, the measures that were used did predict understanding as expected and were correlated with each other, suggesting they are capturing variance in their respective constructs.

The two measures of understanding (Essay Concepts and Inference Tests) similarly correlated with reading skill and with CLEAR Thinking, but correlated only modestly with each other. In addition, the inference test but not the essay concepts were significantly related to prior knowledge. The lack of relation between essays and prior knowledge makes sense given that the documents were available during writing, so students did not need to rely upon retrieval from long term memory in order to construct a more complete argument. Thus, the IVT and essay measures reflect somewhat different aspects of multiple-documents comprehension. Yet, a motivating disposition towards considering evidence (CLEAR Thinking) related to both of these different aspects of comprehension independently from reading skill, prior knowledge, and topic and domain interest. This is consistent with Stanovich's (2012) distinction between individual differences in what a person might be capable of (e.g., reading skill) versus what a person might be disposed to do. This study demonstrates that both are required for successful learning from multiple-document inquiry tasks in science.

Although the current study does not directly test the MD-Trace model, the finding that CLEAR thinking scores predicted middle school students' learning from multiple documents is consistent with the importance of the task model. Bråten et al (2011) hypothesized that epistemic beliefs contribute to the creation of a task model and it is likely that thinking dispositions function in much the same way. It is expected that readers who are disposed to using evidence and reasoning to form and update their beliefs will have a very different task model from those who do not. As a result, they will create different subgoals to guide reading. For example, students with an evidence-based disposition will be expected to seek coherence across explanatory elements and look for evidence to support claims. These subgoals will lead to the integration of more of the key causal concepts from the document set into their mental model of climate change. The present results also highlight that a task model may be more than what a learner thinks is expected and required for learning, but may also include their personal goals related to their desire to learn and update their views versus to protect and maintain their existing views. An interesting direction for future research would be to investigate the manner in which thinking dispositions influence the development of a task model.

Limitations and Future Directions

The working assumption behind the present findings is that students' dispositions towards evidence-based thinking impacts how they approach a multiple-documents inquiry task, Such dispositions make students more likely to engage in the kind of integrative, coherence-building, argumentation processes that have been shown to improve learning in these contexts (e.g., Wiley et al., 2011). The previously reviewed literature shows there is much variance in reading behaviors and strategies when readers are faced with multiple documents. Thinking dispositions may be a generalized individual difference that contributes to this variance. However, we note that the current study only measured the learning outcomes that were presumed to result from these different behaviors, but did not include any on-line measures of processing and reading strategies to verify actual differences in processing. Griffin and Ohlsson (2001) speculated that people who had previously formed a belief on a topic via evidence-based reasoning rather than relying on affective preferences may be better able to represent new belief-relevant concepts. Thus, an alternative to differences in how readers are actively engaging in the task is differences in how their past reasoning on the topic impacts their ability to represent the concepts. Future research is needed to provide evidence that readers vary in their processing during reading in ways that might mediate the observed learning outcomes.

Implications for Instruction

The fact that such a general thinking disposition was able to show relations to learning on a specific topic within science is pedagogically useful. The trend in research on the related construct of epistemology has been toward measuring more domain-specific rather than more general thinking dispositions (e.g., Hofer, 2006). In fact, Braten and Stromso (2010) have even argued for using topic specific epistemology, and have demonstrated that it can be used as a successful predictor of science learning for multiple documents on that topic. Predicting learning does seem to benefit from measuring epistemology in more specific ways (for a review, see Muis, Bendixen, & Haerle, 2006). However, from a pedagogical perspective, identifying general dispositions that could improve learning may be more pragmatically useful as targets for instruction. Domain-specific and topic-specific dispositions imply that separate pedagogies would be needed to target the development of thinking dispositions within each domain or on each specific topic. Any benefit of such interventions would be limited to that domain or topic. The present results suggest that there are more general thinking dispositions regarding the value of evidence that can also have a substantial impact on learning.

There has been little work on interventions targeting an evidence-based disposition. The fact that the disposition itself has some domain generality does not imply that topic-specific learning activities would be ineffective in fostering a dispositional change. Topic specificity may be necessary in order to expose students to examples of such thinking and to have them engage in tasks that require it. Long-term impact may prove difficult if this disposition reflects core values related to commitment to evidence versus the perceived value of sticking to one's beliefs. Such values would seem to be shaped by the social reinforcement students receive in many areas of life outside of school, from their home, religious upbringing, media, and popular culture. On the other hand, there appears to be little existing effort in schools to directly and explicitly foster an evidence-based disposition. Thus, even minor interventions could notably increase students' exposure to the importance and utility of adopting such a disposition. This is another direction for future research. If such a general

disposition can be effectively encouraged, developed or leveraged through instruction, it has the potential to impact learning across topics and domains more generally.

• • •

Acknowledgement

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education in part by Grant R305B07460 Improving Metacomprehension and Self-Regulated Learning from Scientific Texts from the CASL program to Thomas D. Griffin and Jennifer Wiley and in part by Grant R305F100007 Reading for Understanding Across Grades 6 through 12: Evidence-based Argumentation for Disciplinary Learning from the Reading for Understanding Research Initiative to Anne Britt, Thomas Griffin and Jennifer Wiley. The authors thank Karyn Higgs, Kris Kopp, Brent Steffens, Andrew Taylor and the other members of Project READi for their assistance and contributions. The opinions expressed are those of the authors and do not represent views of the Institute of Education Sciences or the U.S. Department of Education.

Thomas D. GRIFFIN. Research Professor, Department of Psychology, University of Illinois at Chicago. His research interests include contexts and individual differences that impact learning in science and history, the impact of affect-based beliefs on the comprehension and acceptance of scientific theories, and the factors that determine metacomprehension monitoring accuracy for explanatory texts.

Jennifer WILEY. Professor, Department of Psychology, University of Illinois at Chicago. Her research interests include contexts that improve learning from informational texts in history and science, and the role of individual differences, especially in working memory capacity and spatial ability in comprehension processes.

M. Anne BRITT. Professor, Department of Psychology, Northern Illinois University. Her research interests include argument comprehension and production, and learning and reasoning from multiple documents.

Carlos R. SALAS. Graduate Student Researcher, Department of Psychology, University of Illinois at Chicago. His research interests include text comprehension, argumentation, and belief formation.

References

Baron, J. (1985). *Rationality and intelligence*. Cambridge, England: Cambridge University Press.

Bennington, J. B. (2009). *The Carbon Cycle and Climate Change*. Cengage Learning.

Bråten, I., Britt, M.A., Strømsø, H.I., & Rouet, J.F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Towards an integrated model. *Educational Psychologist, 46,* 48-70.

Bråten, I., & Strømsø, H.I. (2010). Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse Processes, 47,* 1-31.

Bråten, I., Strømsø, H.I., & Britt, M.A. (2009). Trust matters: Examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly, 44,* 6–28.

Britt, M.A., Perfetti, C.A., Sandak, R., & Rouet, J-F. (1999). Content integration and source separation in learning from multiple texts. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.). *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso (pp. 209-233)*. Mahwah, NJ: Lawrence Erlbaum Associates.

Britt, M.A., & Rouet, J-F. (2012). Learning with multiple documents: Component skills and their acquisition. In M.J. Lawson & J.R. Kirby (Eds.), *The Quality of Learning: Dispositions, Instruction, and Mental Structures*. Cambridge University Press.

Cerdán, R., Vidal-Abarca, E. (2008). The effects of tasks on integrating information from multiple documents. *Journal of Educational Psychology, 100*, 209-222.

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 5, pp. 161-238). Mahwah, NJ: Erlbaum.

Conley, A. M., Pintrich, P. R., Vekiri, J., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology, 29*, 186–204.

Goldman, S .R., Braasch, J. L .G., Wiley, J., Graesser, A. C., & Brodowinska, K. (in press). Comprehending and learning from Internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly*.

Griffin, T. D. (2008). Faith: serving emotional epistemic goals rather than evidence coherence. In V. Sloutsky, B. Love, and K. McRae (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, pp. 2059-2064. Austin, TX: Cognitive Science Society, Inc.

Griffin, T. D., & Ohlsson, S. (2001). Beliefs versus knowledge: A necessary distinction for explaining, predicting and assessing conceptual change 364-369 In J. D. Moore and K. Stenning (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, pp. 364-369. Austin, TX: Cognitive Science Society, Inc.

Griffin, T. D., Wiley, J., & Thiede, K. W. (2008). Individual differences, rereading, and self-explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory & Cognition, 36*, 93-103.

Hofer, B. (2006). Domain specificity of personal epistemology: Resolved questions, persistent issues, new models. *International Journal of Educational Research. 45*, 85-95.

Kintsch, W. (1994). Learning from text. *American Psychologist, 49*, 294–303.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 363-394.

Kokis, J., Macpherson, R., Toplak, M., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology, 83*, 26-52.

Mason, L., Boldrin, A., & Ariasi, N. (*2010*). Epistemic metacognition in context: Evaluating and learning online Information. *Metacognition and Learning, 5*, 67–90.

Muis, K. R., Bendixen, L. D., & Haerle, F. C. (2006). Domain-generality and domainspecificity in personal epistemology research: Philosophical and empirical questions in the development of a theoretical model. *Educational Psychology Review, 18*, 3-54.

Perfetti, C.A., Rouet, J-F., & Britt, M.A. (1999). Towards a theory of documents representation. In H. van Oostendorp & S.R. Goldman (Eds.) *The Construction of Mental Representations during Reading (pp. 99-122)*. Mahwah, NJ: Lawrence Erlbaum Associates.

Rouet, J-F., & Britt, M.A. (2011). Relevance processes in multiple document comprehension. In M.T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text (pp.* 19 - 52). Greenwich, CT: Information Age Publishing.

Sá , W., Kelley, C., Ho, C., & Stanovich, K. E. (2005). Thinking about personal theories: individual differences in the coordination of theory and evidence. *Personality and Individual Differences, 38*, 1149-1161.

Sá, W. C., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology, 91*, 497-510.

Sanchez, C., & Wiley, J. (2006) Effects of working memory capacity on learning from illustrated text. *Memory & Cognition, 34*, 344-355

Sanchez, C., & Wiley, J. (2009). To scroll or not to scroll: Interactions of text presentation and working memory capacity. *Human Factors, 51*, 730-738.

Sanchez, C. A., Wiley, J., & Goldman, S. R. (2006). Teaching students to evaluate source reliability during internet research tasks. *Proceedings of the 7th International Conference of the Learning Sciences*, Bloomington, IN.

Stanovich, K. E. (2012). On the distinction between rationality and intelligence: Implications for understanding individual differences in reasoning. In K. Holyoak & R. Morrison (Eds.) (pp. 343-365), *The Oxford handbook of thinking and reasoning*. New York: Oxford University Press.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology, 89*, 342-357.

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127*, 161-188.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language, 24*, 595–611.

Weinburgh, M. (1995). Gender differences in student attitudes toward science: A meta-analysis of the literature from 1970 to 1991. *Journal of Research in Science Teaching,* 32(4), 387-398.

West, R. F. Stanovich, K. E., & Toplak, M. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions, *Journal of Educational Psychology, 100*, 930-941.

Wiley, J., Ash, I.K., Sanchez, C.A., & Jaeger, A. (2011). Clarifying readers' goals for learning from expository science texts. In M. McCrudden, J. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 353-374). Greenwich, CT: Information Age Publishing.

Wiley, J., Goldman, S., Graesser, A., Sanchez. C., Ash, I., & Hemmerich, J. (2009). Source evaluation, comprehension, and learning in internet science inquiry tasks. *American Educational Research Journal, 46*, 1060-1106.

Wiley, J., Griffin, T. D., & Thiede, K. W. (2005). Putting the comprehension in meta-comprehension. *Journal of General Psychology, 132*, 408-428.

Wiley, J., & Voss, J. F. (1996). The effects of "playing" historian on learning in history. *Applied Cognitive Psychology, 10*, 63-72.

Wiley, J., & Voss, J. F. (1999) Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology, 91*, 301-311.