

DISRUPTIVE PROMISE: THE LINKS INTERVENTION AND ITS IMPACT ON  
MULTIPLE, MULTIMODAL, INTERNET TEXT INTEGRATION

By

Michelle Schira Hagerman

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

Educational Psychology and Educational Technology—Doctor of Philosophy

2014

## ABSTRACT

### DISRUPTIVE PROMISE: THE LINKS INTERVENTION AND ITS IMPACT ON MULTIPLE, MULTIMODAL INTERNET TEXT INTEGRATION

By

Michelle Schira Hagerman

This dissertation study presents an instructional intervention called LINKS: *Learning to Integrate InterNet Knowledge Strategically*. It reports evidence of the intervention's impact on two variables: (a) ninth graders' use of ten online reading and integration strategies while engaged in dyadic online inquiry on science topics in school, and (b) evidence of integration from multiple, multimodal Internet texts in their written persuasive arguments. Dyads ( $n = 8$ ) were randomly assigned to treatment and control conditions. They completed a pretest, three practice/intervention sessions and a posttest. Groups were matched on pretest reading comprehension scores. The treatment group received LINKS, which included explicit instruction of strategies, modeling of strategy use during think-aloud screencasts, and guided instructional support that prompted students to engage strategies while reading. Teacher support was gradually released over three intervention sessions that lasted approximately one hour each. The control group did not receive instruction on strategies. They watched screencasts that included the same content as treatment screencasts, but received no modeled think aloud of strategy use. Control group participants received no instructional support during their online inquiry sessions. All participants read texts online in dyads but wrote persuasive arguments independently. When assumptions for parametric tests were met, they were used for between-groups and within-groups comparisons. Usually non-parametric Mann Whitney-U tests were used for between-groups comparisons and Wilcoxon Signed Rank tests were used for within-groups comparisons. During

intervention sessions, treatment participants were more likely than control participants to focus on the type of text and the trustworthiness of a text as they considered its relevance or utility. At posttest, however, treatment condition participants performed no differently from the control participants in terms of these critical evaluation strategies. At posttest, treatment participants were more likely than control participants to explicitly engage background knowledge during online inquiry. This was the only significant posttest difference in strategy use between the groups. In their written arguments, treatment participants were also more likely to use facts they had noted or recorded as preexisting knowledge during inquiry. Although index scores on the Trace Indicators of Integration (TII) Rubric did not differ statistically at pretest or at posttest for the groups, treatment participants did see a statistically significant improvement in their TII index scores by the second intervention session. A similar bump in performance was not observed for the control condition. At posttest, treatment participants were also more likely to include in their written arguments counter points gathered from websites that differed from those used to build the main argument, suggesting that LINKS may have enabled this group of students to bring together more perspectives from a broader range of Internet texts in their written arguments. A single case analysis suggests that students who are very early in their learning trajectories for multiple, multimodal Internet text integration skills may benefit considerably from LINKS. Methodologically, this dissertation also introduces a protocol for measuring trace evidence of integration in students' written arguments. Although results should be interpreted cautiously, teachers of adolescents may find that LINKS offers a promising place to start instruction for online inquiry and the construction of meaning across multiple, multimodal Internet texts.

Copyright by  
MICHELLE SCHIRA HAGERMAN  
2014

## ACKNOWLEDGEMENTS

In ways big, small, personal, public, long ago and very recently, many people and organizations have enabled me to do this work. Without their influence, trust, and support I may not have thought it possible in the first place, had the means to make it happen, known how to get it done, or known why to do it at all.

I extend thanks to my advisor, Dr. Douglas K. Hartman for apprenticing me into the literacies research community, and for helping me to refine my ideas. I also thank my committee, Dr. Rand J. Spiro, Dr. Kelly Mix and Dr. Nell K. Duke for their thought leadership and guidance.

The students who participated in this research, the teachers and principals who allowed them to be released from their regular class activities, and the families who permitted their children to participate deserve special acknowledgement. My research depended on your spirit of volunteerism and in the trust you put in me and in this research. Thank you for these gifts.

Faculty of the Educational Psychology and Educational Technology program at MSU gave me entry to this work, and have each supported my development as a scholar along the way. Thank you especially to Dr. Leigh Graves Wolf, Dr. Punya Mishra, and Dr. Richard Prawat for your mentorship and investment in my development as a scholar and teacher. I also thank the Graduate School at Michigan State University and the Social Sciences and Humanities Research Council of Canada for funding that enabled me to develop as a researcher.

Erricka Blosser of Albion College deserves special thanks for her assistance through the coding and analysis of these data. I thank Angie Johnson, Paul Morsink, and Selena Protacio Damian for helping me to improve this work along the way. These three colleagues provided substantive expertise that strengthened my arguments and clarified my thinking. Michael

DeSchryver and Erica Hamilton offered thoughtful suggestions that pushed me to improve this work when I really needed the help. Amber White of North Branch Areas Schools deserves special thanks for trusting in my ideas and for giving me new perspectives on their potential to support teachers and students in classrooms. Julie Smith—merci.

My friends and family deserve the deepest and most heartfelt thanks of all. My parents provided the foundation on which I have grown. I thank them, and my parents-in-law too, for their infinite gifts of love, time and encouragement. My friends have boosted my spirits, cheered me on, and inspired me to persist. Thank you.

My biggest thanks, though, are reserved for the three people who have given the most and without whom this never would have been possible. Zoë, Ellyn and Chris -- in ways big, small, personal, public, long ago and very recently, you have given me everything so that I could do this. Je vous aime infiniment. AΩ.

## TABLE OF CONTENTS

LIST OF TABLES.....	x
LIST OF FIGURES .....	xii
INTRODUCTION .....	1
<b>Key Terms</b> .....	4
CHAPTER 1: Theoretical Framework .....	7
<b>Documents Model of Multiple Text Integration</b> .....	7
<b>Document Search Processes</b> .....	8
<b>Schemas of the Moment</b> .....	9
<b>New Literacies</b> .....	9
<b>Writing to Learn</b> .....	10
<b>Socio-cultural and Situated Perspectives on Learning</b> .....	10
CHAPTER 2: Review Of Literature.....	12
<b>What To Teach</b> .....	12
<b>[(PST)<sup>2</sup> + (iC<sup>3</sup>)]</b> .....	17
<b>How To Teach [(PST)<sup>2</sup> + (iC<sup>3</sup>)] To Ninth Grade Students</b> .....	18
<b>LINKS</b> .....	23
<b>Research Questions</b> .....	23
CHAPTER 3: Method.....	25
<b>Design</b> .....	25
<b>Participants</b> .....	27
<b>Research Context</b> .....	31
<b>Assignment to Dyad</b> .....	33
<b>Sequence and Pacing of Research Sessions</b> .....	33
<b>Instruments</b> .....	34
<b>Reading passages subtest of the Woodcock Reading Mastery Test</b> .....	34
<b>Recorded audio, video and navigational clickstream data</b> .....	35
<b>Recorded notes and background knowledge</b> .....	35
<b>Written persuasive arguments</b> .....	36
<b><i>ACT persuasive argument scoring guidelines</i></b> .....	36
<b><i>Trace indicators of integration rubric</i></b> .....	37
<b>Conditions of Treatment</b> .....	38
<b>Treatment condition</b> .....	38
<b>Control condition</b> .....	39
<b>Screencasts</b> .....	39
<b>Guided Questioning</b> .....	41
<b>Treatment condition</b> .....	41
<b>Control condition</b> .....	43
<b>Topic Prompts</b> .....	44
<b>Data Analyses</b> .....	45

<b>Data analyses for research question 1</b> .....	45
<i>Coding manual</i> .....	45
<i>Unit of analysis</i> .....	47
<i>Interrater agreement</i> .....	50
<i>Phase 1</i> .....	50
<i>Phase 2</i> .....	51
<i>Frequency counts</i> .....	51
<i>Relative frequencies</i> .....	51
<i>Relative duration</i> .....	51
<i>Mean gap</i> .....	52
<b>Data analyses for research question 2</b> .....	52
<i>Evaluation of persuasive arguments for overall quality</i> .....	52
<i>Evaluation of persuasive arguments for evidence of integration</i> .....	52
<b>CHAPTER 4: Results</b> .....	55
<b>Section 1: LINKS Impact on Strategy Application</b> .....	55
<b>Pretest comparison of reading scores</b> .....	55
<b>Comparisons of mean frequencies of total strategy use</b> .....	56
<i>Comparison of ranges of difference scores of total strategy use</i> .....	59
<b>Frequencies comparisons by strategy</b> .....	60
<i>Descriptive analyses of strategy by strategy frequencies data</i> .....	62
<i>Strategy by strategy frequencies comparisons</i> .....	64
<i>Comparisons of strategy by strategy relative frequencies</i> .....	66
<i>Within-groups analysis of change in Type (E)</i> .....	68
<i>Within-groups analysis of change in Trustworthiness (F)</i> .....	70
<b>Analyses of time patterns</b> .....	72
<i>Descriptive analysis of time patterns</i> .....	73
<i>Comparisons of relative duration</i> .....	73
<i>Comparisons of mean gap</i> .....	77
<b>Summary of section 1 results</b> .....	78
<b>Section 2: LINKS' Impact on Trace Indicators of Integration in Students' Writing</b> .....	80
<b>Evaluation of pretest quality of written persuasive arguments</b> .....	80
<b>Comparison of trace indicators of integration in persuasive arguments</b> .....	81
<b>Comparison of discrete indicators of integration</b> .....	83
<i>Integration in students' written persuasive arguments</i> .....	86
<i>Example 1: Higher scoring essay [16/20 on TII index][Control participant, Practice Session 1]</i> .....	87
<i>Example 2: Lower scoring essay [8/20 on TII index][Treatment participant, Practice Session 1]</i> .....	89
<b>Summary of section 2 results</b> .....	91
<b>Section 3: Case Study Analysis Of Dyad 8</b> .....	92
<b>Evidence of change in writing</b> .....	93
<i>Pretest persuasive arguments</i> .....	93
<i>John's pretest written persuasive argument [1/20 on TII index]</i> .....	93
<i>Alex's pretest written persuasive argument [3/20 on TII index]</i> .....	94
<i>Posttest persuasive arguments</i> .....	95



<i>John's posttest written persuasive argument [7/20 on TII index]</i> .....	95
<i>Alex's posttest written persuasive argument [10/20 on TII index]</i> .....	96
<b>Evidence of change during online inquiry</b> .....	98
<i>Relative duration of strategies use</i> .....	99
<i>Mean gap of strategies use</i> .....	102
<b>Qualitative comparison of integration in Dyad 8's thinking</b> .....	103
<b>Summary of section 3 results</b> .....	111
<b>CHAPTER 5: General Discussion</b> .....	114
<b>Research Question 1: Impact On Strategy Application</b> .....	115
<b>Research Question 2: Impact On Trace Evidence Of Integration In Persuasive Arguments</b> .....	118
<b>Why Was LINKS Disruptive?</b> .....	120
<b>Questions requiring further analysis</b> .....	124
<b>Limitations</b> .....	126
<b>CHAPTER 6: Conclusion</b> .....	129
<b>APPENDICES</b> .....	132
Appendix A: ACT Persuasive Writing Scoring Guidelines .....	133
Appendix B: Trace Indicators of Integration (TII) Index Rubric .....	136
Appendix C: Screencast Transcripts .....	138
Appendix D: Screencast URLs and Timings .....	150
Appendix E: Topic Prompts .....	152
Appendix F: List of Codes .....	154
Appendix G: Representative Excerpts for Each Coding Category .....	158
Appendix H: Representative Essay Evaluations and Markups .....	169
Appendix I: John and Alex's Evaluations .....	175
<b>REFERENCES</b> .....	183

## LIST OF TABLES

Table 1 <i>Summary of the <math>[(PST)^2 + (iC^3)]</math> Framework of Multiple, Multimodal Internet Text Integration Strategies.....</i>	19
Table 2 <i>Summary of Study Design and Activities for Treatment and Control Groups.....</i>	26
Table 3 <i>Summary of Participation by School.....</i>	29
Table 4 <i>Summary of Descriptive Data for Control and Treatment Dyads.....</i>	30
Table 5 <i>Summary of Screencast Times .....</i>	41
Table 6 <i>Summary of Letter Label and Strategy Correspondence .....</i>	46
Table 7 <i>Mean Frequencies of Total Strategic Episodes.....</i>	57
Table 8 <i>Difference Scores for all Dyads.....</i>	60
Table 9 <i>Pretest and Posttest Mean Strategic Episodes <math>[(PST)^2 + (iC^3)]</math> for Control Group.....</i>	61
Table 10 <i>Pretest and Posttest Mean Strategic Episodes <math>[(PST)^2 + (iC^3)]</math> for Treatment Group... </i>	61
Table 11 <i>Comparison of Pretest and Posttest Mean Relative Frequencies <math>[(PST)^2 + (iC^3)]</math> for Control Group.....</i>	67
Table 12 <i>Comparison of Pretest and Posttest Mean Relative Frequencies <math>[(PST)^2 + (iC^3)]</math> for Treatment Group.....</i>	68
Table 13 <i>Mean Relative Frequencies of Type (E) Episodes by Condition .....</i>	69
Table 14 <i>Mean Relative Frequencies of Trustworthiness (F) Episodes by Condition.....</i>	71
Table 15 <i>Mean Values of Relative Duration for Control and Treatment Groups.....</i>	77
Table 16 <i>Summary of Mean TII Scores for Control and Treatment Groups.....</i>	81
Table D1 <i>Summary of Screencast URLs.....</i>	149
Table D2 <i>Summary of Screencast Content and Timings.....</i>	150
Table H1 <i>TII Rubric Evaluation of Higher Scoring Essay.....</i>	170
Table H2 <i>TII Rubric Evaluation of Lower Scoring Essay.....</i>	173

Table I1 <i>John's Pretest TII Evaluation [Score = 1/20]</i> .....	175
Table I2 <i>Alex's Pretest TII Evaluation [Score = 3/20]</i> .....	177
Table I3 <i>John's Posttest TII Evaluation [Score = 7/20]</i> .....	179
Table I4 <i>Alex's Posttest TII Evaluation [Score = 10/20]</i> .....	181

## LIST OF FIGURES

<i>Figure 1.</i> Total strategic processing episodes over five sessions for control dyads (n=4) .....	58
<i>Figure 2.</i> Total strategic processing episodes over five sessions for treatment dyads (n=4) .....	59
<i>Figure 3.</i> Pyramid representation of search, source selection, and identify important information (n=8) with mean frequency (M) and [95% confidence intervals] reported .....	62
<i>Figure 4.</i> Mean relative frequency of Type (E) episodes over five sessions of the study for treatment and control groups .....	70
<i>Figure 5.</i> Mean relative frequency of Trustworthiness (F) episodes over five sessions of the study for treatment and control groups .....	71
<i>Figure 6.</i> Control group: Pretest relative duration.....	75
<i>Figure 7.</i> Control group: Posttest relative duration.....	75
<i>Figure 8.</i> Treatment group: Pretest relative duration.....	75
<i>Figure 9.</i> Treatment group: Posttest relative duration .....	75
<i>Figure 10.</i> Control group: Pretest mean gap.....	76
<i>Figure 11.</i> Control group: Posttest mean gap.....	76
<i>Figure 12.</i> Treatment group: Pretest mean gap.....	76
<i>Figure 13.</i> Treatment group: Posttest mean gap.....	76
<i>Figure 14.</i> Mean Trace Indicators of Integration Index scores for both groups on all five essays .....	83
<i>Figure 15.</i> Relative duration of strategy application for Dyad 8 over time.....	102
<i>Figure 16.</i> Mean gap for of Strategy Application for Dyad 8. ....	103

## INTRODUCTION

The Common Core State Standards for English Language Arts, History/Social Studies, Science and the Technical Subjects (CCSS) (National Governors' Association Center for Best Practices & Council of Chief School Officers, 2010) expect K-12 students to gather, critically evaluate, and integrate knowledge and ideas from multiple print and digital sources.<sup>1</sup> This expectation is articulated in CCSS anchor standards for reading, writing, listening and speaking. At every grade level, in every subject, and as a part of every strand of literacies instruction outlined in the CCSS, this expectation for students to integrate knowledge and ideas is framing teachers' focus, and students' learning, in American schools. Though the design and merits of this new curriculum have been debated (e.g., Drew, 2013; Hiebert & Mesmer, 2013), on the particular expectation for integration of knowledge and ideas from multiple texts, there is good reason to invest in research that documents and clarifies both how multiple text integration happens on the Internet, and how to support students' ability to construct an integrated understanding from what they read across multiple digital texts.

Rouet (2006) has described document integration and information search as “compulsory elements of *functional literacy*, especially in a world that relies more and more strongly on sophisticated digital information systems” (p. 189). To him, multiple text integration skills, particularly on the Internet, are absolute requirements for life in a digital world. DeSchryver

---

<sup>1</sup> According to the Common Core State Standards website on August 16, 2013, the standards have been adopted by 45 US states, the District of Columbia, US Department of Defense Schools, Guam, American Samoa and the US Virgin Islands (<http://www.corestandards.org/in-the-states>)

(2012) agrees. He argues that just knowing how to find and understand information on the Web will not enable the kind of complex thinking required to address the most serious issues of our times (p. 4). Integration and, as he notes, the generation of new understanding from the juxtaposition of multiple online texts, are among the advanced literacies skills now required in “an age of complexity” (p. 4). The imperative for teaching multiple Internet text integration skills is therefore driven by curriculum expectations in schools, but even more importantly, by these broader understandings that very advanced digital literacies skills are vital to the literate lives of all students.

Many studies of multiple text integration with printed sources have shown, however, that students generally struggle to construct an integrated understanding of curricular topics from multiple texts (Britt & Aglinskias, 2002; Mateos, Martín, Villalón & Luna, 2008; Spivey & King, 1989; Stahl, Hynd, Britton, McNish & Bosquet, 1996; Wineburg, 1991; Wolfe & Goldman, 2005). Unlike experts who leverage disciplinary knowledge of texts (Moje, Stockdill, Kim & Kim, 2011; Shanahan, 2009) and heuristics such as corroboration, sourcing and contextualization (Rouet, 2006; Wineburg, 1991), novice integrators prioritize content relevance over source reliability, (Wineburg, 1991) overlook contrasting points of view in texts (Britt & Aglinskias, 2002) and may be less aware of meaning-laden structural cues in texts that can support understanding (Goldman & Rakestraw, 2000; Wolfe & Goldman, 2005).

Online, studies have begun to reveal similar discrepancies between more and less expert integrators, particularly when it comes to critical evaluation of texts (Goldman, Braasch, Wiley, Graesser & Brodowinska, 2012; Wiley, Goldman, Graesser, Sanchez, Ash, & Hemmerich, 2009). Relative to their less expert peers, more capable online readers spend more time on more reliable Internet texts, engage in more comprehension-monitoring processes at these reliable sites

and make navigation decisions that are more goal-oriented (Goldman et al., 2012). Evidence also suggests that better online readers are more metacognitive. They seem to purposefully and flexibly apply a range of strategies that are unique to the online reading context (Afflerbach & Cho, 2009; Coiro & Dobler, 2007; Sevensma, 2013; Small, Moody, Siddarth & Bookheimer, 2009) and also generate more extra-textual connections and creative syntheses from what they read (DeSchryver, 2012).

For teachers, especially those working in US schools to implement the CCSS expectations for integration of knowledge and ideas, these descriptive findings of more and less expert multiple-texts reading processes, particularly on the Internet, lead to two important questions: First, what should teachers teach to students that will move them toward more expert integration of multiple Internet texts, and second, how should teachers teach that content? At present, the research base is too slim to inform comprehensive answers.

Although a few studies have examined interventions that support online reading comprehension processes generally (e.g., Castek, 2008; Colwell, Hunt-Barron & Reinking, 2013; Dwyer, 2010) and critical evaluation, more specifically (Braasch, Bråten, Strømsø, Anmarkrud & Ferguson, 2013; Macedo-Rouet, Braasch, Britt & Rouet, 2013; Wiley et al., 2009; Zhang & Duke, 2011) it is not yet clear which methods, bundled together with which strategies of instructional focus, might enable students to construct the most integrated understandings of the texts they read online. Importantly, these questions are urgently needed to inform instruction for students who may be just learning to integrate multiple Internet texts for their school-related inquiry projects.

Given the need for this research, I used theory and descriptive data to design a set of ten multiple text integration strategies that I hypothesized, if applied flexibly and recursively, would

support multiple text integration processes and lead to a more integrated understanding of science topics. Over three practice sessions, I taught the ten strategies to ninth-graders using direct instruction, think-aloud modeling recorded in a series of screencasts, and guided questioning during reading and research sessions. In this study, I present evidence of the intervention's impact on (a) strategy application during Internet reading and research, and (b) indicators of integration in persuasive essays that participants wrote immediately after their reading and research sessions.

### **Key Terms**

This is a study about ninth-graders reading and learning to integrate meaning from multiple, multimodal Internet texts. As such, the terms *integration*, *texts*, and *multimodal* warrant precise definition. I also define the related terms *document*, *source*, and *online inquiry* because they are important to the research.

Integration: Also called synthesis (Leu, Kinzer, Coiro & Cammack, 2004), integration is the recursive, iterative process of cognitive bricolage that leads to the construction of a Documents Model of understanding (Britt, Rouet & Braasch, 2013; Perfetti, Rouet & Britt, 1999; Rouet, 2006) from multiple texts. For the purposes of this study, the texts are Internet texts. Integration is the putting together of meaning from multiple texts, found and read on the Internet. [The Documents Model is described in the Theoretical Framework.]

Online inquiry: Also called online reading comprehension (Coiro, 2011a, 2011b ; Leu, Kinzer, Coiro & Cammack, 2004) *online inquiry* refers to the set of processes involved in the research of a topic, using the Internet. It requires a focus on the purpose(s) of inquiry, the iterative generation of questions that lead to answers, the location of information using search engines, the critical evaluation of information resources (Goldman & Scardamalia, 2013), the



extraction of relevant and trustworthy information from each information resource, integration of that information (see definition of integration, above) and the communication of understanding for specific purpose(s) and to specific audience(s).

Text: A bounded entity of meaning, of variable structure (e.g., webpage, blog post, forum comment) or modality (e.g., words, video, audio, graphic, computational formula, color, white space) found on the Internet, and used to inform understanding of a topic or question. This definition of *text* follows from Fox and Alexander's (2009) descriptions of "transitional extensions" (p. 233) of more traditional definitions of texts and text comprehension activities which have been grounded in printed pages, paragraphs, sentences and words. On the Internet, texts can be "fluid or static" in structure and, "in single or multiple modalities of single or multiple linked propositional networks" (Fox & Alexander, 2009, p. 223). Web-based documents often include more than one text or type of text (Britt, et al., 2013).

Document: A text or set of texts of interest or relevance to the reader due to its features of authorship, context, informational content and/or rhetorical purpose (Britt, Rouet and Braasch, 2013). Documents include source information that enable the reader to critically evaluate its value.

Source: Attributional information such as authorship, authorship credentials, context of publication, date of publication, or authorship purpose that enables readers to critically evaluate the trustworthiness of a text and position it relative to others. This definition follows from Wineburg's (1991) definition of sourcing as the act of looking to the source before reading a text closely (p. 77). The term *source selection*, one of the ten strategies modeled for participants in this research, was chosen (rather than documents selection, for instance) as a way of placing emphasis on attributional information, and the evaluation of text and document trustworthiness.

Multimodal: The description of texts and documents that include “more than one mode of meaning making” (Kress, Jewitt, Ogborn & Tsatsareli, 2001, p. 42). Modes are assumed to include semiotic representations of meaning that engage visual, acoustic and spatial processes along with those processes required to construct meaning from words. Video, music, podcasts, infographics, graphs, charts, rating scales, color, use of white space, use of text features such as boldface or italics -- any of these modes (and more) constitute important elements of meaning that the online reader can use to construct understanding of texts, documents and sources (Kress, 2010; Kress & van Leeuwen, 2001; Roswell & Burke, 2009). Following from Jewitt (2008) I broadly assume that the "form of representation is integral to meaning and learning more generally" (p. 241).

## CHAPTER 1

### **Theoretical Framework**

Before turning to a more thorough review of the literature that informed the design of the instructional intervention and its implementation, it is important to identify the larger, interrelated theoretical assumptions that undergird this work.

#### **Documents Model of Multiple Text Integration**

In this study, I assume that processes of multiple text integration are the cognitive processes and strategies that lead to the development of an integrated mental representation, or documents model (Britt, et al., 2013; Perfetti, et al., 1999; Rouet, 2006) of the topic in question. I assume that this representation takes the form of schemas (Anderson & Pearson, 1984) that constitute both the substrata and the result of understanding. In particular, I borrow from the Construction-Integration Model of reading comprehension (Kintsch, 1998; Kintsch & van Dijk, 1978) and The Documents Model of multiple text integration (Britt, et al., 2013; Perfetti, et al., 1999; Rouet, 2006;) as footings for the intervention's design.

The construction-integration model of reading comprehension (Kintsch, 1998; Kintsch & van Dijk, 1978) posits that readers construct a model of understanding within a single text by first building a text base and then a situation model for the text. Broadly speaking, readers decode to build the text base. They also use foundational linguistic knowledge such as syntax and semantics to understand the basic meanings of idea units (propositions) in the text. Of course, no two readers will interpret texts in exactly the same way and the situation model accounts for this. The situation model is the reader's understanding of the text base that integrates background knowledge. The full mental model of understanding that readers construct,

is therefore informed by foundational reading processes, and reading knowledge that is linguistic and experiential.

Building from the construction-integration model (Kintsch, 1998; Kintsch & van Dijk, 1978) Perfetti, et al. (1999) suggest that the mental models that expert readers develop when they read and integrate understanding of multiple texts involves “mental representations of specific texts, situations described in texts, and relations among texts” (p. 99). Rouet (2006) further contends that knowledge of source and content and the way that these two factors connect to one another, permits single documents to be synthesized into an integrated, multiple documents model of understanding (pp. 71-72). According to this view, source information gives readers a framework for comparing content; it is a tool that permits relative weighting when information from many sources differs (p. 74). I therefore assume that an instructional focus on content, source and the relationships among documents may move novice online readers toward a more expert level of multiple texts integration as they read online.

### **Document Search Processes**

In this study, I assume that online inquiry will depend, in part, on the ability to purposefully search the Internet, but to also find and extract information that contributes to the development of an argument, as outlined by the task prompt. Rouet’s Task-Based Relevance Assessment and Content Extraction (TRACE) model (2006, p. 105) extends to print-based and Internet contexts, and its core assumptions were used to inform the design of the intervention. According to TRACE, the central components of text processing are (a) the construction of a task model based on internal needs and environmental constraints, (b) the assessment of document relevance, based on available information resources and search tools, (c) the cyclical extraction and integration of content information in order to construct an internal response model

(Rouet, 2006, p. 105). This research assumes participants engage in the TRACE process to construct an integrated mental model of understanding.

### **Schemas of the Moment**

Whereas Rouet's TRACE model provides a general model for thinking about the cognitive processes engaged for purpose-driven online search or inquiry, Spiro and Deschryver (2009) suggest that the iterative search cycle, which can occur very quickly on the Internet, may result in the juxtaposition of multiple perspectives and many alternative points of connection in ways that are unique for each reader. This rapid "criss-crossing of the Web landscape" means that readers explore "many potential situation-sensitive knowledge assembly paths to build *schemas of the moment* to suit the needs of unforeseeable future situations" (p. 116). Consistent with this view, in the development of this intervention, I assumed that multiple text integration online requires and is influenced by the reader's own "path construction". Moreover, I assume that rapid criss-crossing is unique to the Internet context, promotes the juxtaposition of multiple perspectives and the construction of schemas of the moment that, ultimately, contribute to the reader's integrated model of understanding (Perfetti et al., 1999; Rouet, 2006).

### **New Literacies**

The theory of New Literacies (Coiro, Knobel, Lankshear & Leu, 2008; Leu, et al., 2004) contends that when readers construct meaning from Internet texts, the skills and strategies that they use with printed texts (e.g., Pressley & Afflerbach, 1995) are essential, but insufficient (Coiro, 2011a). According to this view, the particularities of the Internet, with its unique culture, text genres, text structures, navigational demands, authorship practices, and interactive affordances require readers to adopt skills, strategies and mindsets that include, but also extend beyond those required to construct meaning from printed texts (Afflerbach & Cho, 2009; Coiro

& Dobler, 2007; Hartman, Morsink & Zheng, 2010). Specifically, this theory suggests that online reading comprehension is a process of online inquiry, driven by purpose (Coiro, 2011b; Coiro & Hagerman, 2013) that requires the application of new ways of questioning, locating, evaluating, synthesizing and, communicating information (Leu, et al., 2004). The theory also contends that these skills can, and indeed should, be taught in schools (Coiro, 2011a, 2011b; Henry, 2006; Leu, 2000; Leu, Coiro, Castek, Hartman, Henry & Reinking, 2008; Leu, D. J., et al., 2012a, 2012b). The central tenets of this theory framed the design of the intervention.

### **Writing to Learn**

Although writing is not the focus of the intervention, theories of writing to learn informed its design. This theoretical view posits that writing, particularly in the disciplines, supports the construction of meaning (e.g., Klein & Rose, 2010; Langer, 1986a, 1986b; Newell, 2006). I especially leverage Langer's foundational finding that when students write essays, "they seem to step back from the text after reading it – they reconceptualize the content in ways that cut across ideas, focusing on larger issues or topics. In doing this, they integrate information and engage in more complex thought" (1986a, p. 406). From this research, I assumed that the act of writing itself would support multiple-text integration, and that trace evidence of integration processes would be evident in participants' written arguments. Moreover, I assume that the organization and the content included in students' written arguments represent students' understanding of the topic.

### **Socio-cultural and Situated Perspectives on Learning**

The intervention is also grounded in the theoretical view that learning and the context in which the learning happens are inextricably connected (Brown, Collins & Duguid, 1989; Lave & Wenger, 1991, Vygotsky, 1978). It is informed by the related theoretical understanding that

learning contexts are shaped by social systems and that children, indeed, all novices, learn through a model of social apprenticeship that is responsive to the learner's changing needs. The design of the study and intervention engage social supports that reflect socio-cultural, and situated views of learning. The study took place in students' schools, the context where children are expected to develop multiple text integration skills. All students read with a partner in order to leverage the social, discursive nature of knowledge construction with texts that has been described by Bakhtin (1981) and Hartman (1995). It was assumed that readers "borrow and link the texts of others in constructing their own inner texts" (Hartman, 1995, p. 530). The treatment scaffolds (Duke & Pearson, 2002; Duke, Pearson, Strachan & Billman, 2011; Pearson & Gallagher, 1983) students' acquisition of multiple text integration skills by providing direct instruction in the learning medium (i.e., video), but also teacher questioning and guidance that was responsive to students' learning needs (Vygotsky, 1978).

## CHAPTER 2

### Review Of Literature

In the first part of this literature review, I summarize findings that juxtapose both the promise and the challenges of constructing an integrated mental model of understanding from multiple texts. These findings informed *the what* of the intervention, the  $[(PST)^2 + (iC^3)]$  framework. I describe  $[(PST)^2 + (iC^3)]$  at the end of this section.

The second part of the review summarizes pedagogical practices and interventions that informed my choices of instructional design. These studies informed *the how* of the intervention, which I've called LINKS [Learning to Integrate InterNet Knowledge Strategically]. I describe the choices I made for how to teach  $[(PST)^2 + (iC^3)]$ , through LINKS at the end of this section.

#### What to Teach

It is generally understood that whatever the text, whether online or offline, expert reading is varied, strategic, flexible, informed by prior knowledge and driven by awareness of purpose (Afflerbach & Cho, 2009; Coiro & Dobler, 2007; Pressley & Afflerbach, 1995; RAND Reading Study Group, 2002; Spiro, Coulson, Feltovitch & Anderson, 2004; Zhang & Duke, 2008). Shanahan (2009) notes, "Competent readers do not use a universal approach to reading. Depending on the level of prior knowledge, the kind of text, and the purpose for reading, individuals alter their attention to different structural, rhetorical, and linguistic characteristics and think in varied ways about the elements they encounter." (p. 240). Ideally, the teaching of reading in any context should therefore support development of complex, strategic and flexible approaches to meaning making.

Evidence from studies of expert multiple text integration elucidate processes and strategies that can inform the question of what to teach so that students become better multiple



text integrators. As already noted, offline and online, capable integrators navigate a course of meaning construction that includes building a text base and a situation model (Kintsch, 1998; Kintsch & vanDijk, 1978; McNamara & Shapiro, 2005) for each text. Next, they identify and build connections among texts (Perfetti et al., 1999; Rouet, 2006). To do this, good readers of multiple Internet texts evaluate potential texts for content relevance (Rouet, 2006; Wiley et al., 2009). This is often done in the context of the search engine results page (SERP) where, as Coiro and Dobler (2007) noted, proficient online readers are very aware of their processes of forward inferencing before clicking on a link. Afflerbach and Cho (2009) describe initial evaluation of content utility or relevance as one strategy for “realizing and constructing potential texts to read” (p. 82). They also note that good readers “sample goal-related information at the initial stage of reading to establish a dynamic plan to achieve one’s own goal” (p. 82). Teaching students how to identify and determine relevance based on reading purpose, and how to make inferences about the relevance of texts from cues at the SERP are therefore important skills for students to learn.

Good online readers also evaluate trustworthiness (Braasch et al., 2013; Bråten, Strømsø & Britt, 2009; Goldman, et al., 2012; Wiley et al., 2009) by using sourcing cues such as authorship, (Wineburg, 1991; Rouet, 2006) snippet text, and clues in the URL. They also seem to leverage signals of trustworthiness from text structure and aesthetic design (e.g., Lindgaard, Dudek, Sen, Sumegi, & Noonan, 2011; Wang & Emurian, 2005), text genre (e.g., e-commerce site, blog, forum, video), its’ intended audience, purpose, tone and feel (Afflerbach & Cho, 2009). Teaching students to evaluate trustworthiness using a range of cues also seems important.

McNamara and Shapiro (2005) note that the construction of a cohesive situation model from multiple linked hypertexts is dependent on the structure of the hypertext environment itself, but also on the reader’s pre-existing domain knowledge. Readers with more content knowledge

are more able to construct meaning in open hypertext systems whereas readers with less content knowledge benefit from hypertext environments that explicitly cue the relationships among texts. To support comprehension and by extension, integration of meaning across hypertexts as students explore the open web, this research suggests students could benefit from knowing something about the topic before they begin to read online.

Offline, good multiple text integrators also corroborate facts, looking for similarities and differences among the texts they read (Rouet, Favart, Britt & Perfetti, 1997; Stahl, Hynd, Britton, McNish & Bosquet, 1996; Wineburg, 1991). After reading texts closely, and extracting salient content, good readers weigh the relative value of the information they've gathered to construct an overarching mental model of understanding that includes multiple ideas (Cerdán & Vidal-Abarca, 2008; Kintsch, 1998; Rouet, 2006). This suggests that lessons focused on multiple text integration should teach students to compare, contrast and connect the information they have gathered.

Studies that have compared stronger and weaker online readers can also clarify the skills that should be taught as part of an effective instructional intervention focused on multiple text integration. Broadly speaking, weaker readers have been found to struggle with some or all of the strategies outlined above (Bazerman, 1985; Cerdán & Vidal-Abarca, 2008; Goldman et al., 2012; Sevensma, 2013; Wiley et al., 2009; Wineburg, 1991). Self-regulatory skills that enable readers to focus on purpose, find relevant content, and minimize cognitive load (Afflerbach & Cho, 2009; Azevedo & Cromley, 2004; Balcytienne, 1999; Bråten & Strømsø, 2011; Dwyer, 2010; Eveland & Dunwoody, 2000; Goldman et al., 2012; Sevensma, 2013) seem especially relevant, and have been found to differentiate students who are more and less able to construct an understanding of what they have read online.

For instance, Sevensma (2013) found that at-risk ninth-grade readers were more likely than their typically developing peers to follow “ineffective traversals” (p. 201), navigational paths that returned no understanding or led to confusion as they researched a science topic online. Sevensma also found these students spent time reading on topics that were irrelevant to their research purpose and provided ineffective or very limited “declarative, procedural, and conditional knowledge about strategies that would support their comprehension of the texts” (p. 209). It would seem, then, that methods that reduce cognitive load, support active self-regulation, enable students to become aware of the strategies that would most effectively support their reading processes, and allow them to practice articulating those strategies could be helpful.

In their study of better and poorer undergraduate learners, Goldman et al. (2012) also found discrepancies in strategy use. Poorer learners used fewer monitoring and evaluation strategies than better learners (p. 375). Better learners’ stated reasons for leaving websites also reflected “greater planfulness and goal-directedness” (p. 370).

Emerging evidence also suggests that undergraduate students who are more able to evaluate the trustworthiness of texts, also learn more content-related concepts from their online research. Wiley et al. (2009) found that when their college-age students were instructed to SEEK: (a) think about the source of the information, (b) consider the nature of the evidence in the text, (c) analyze the fit of evidence with an explanation, and (d) compare the fit of the new information with prior knowledge (SEEK) (p.1087) their understanding of scientific phenomena, constructed across multiple texts that they located on the Internet, increased (p. 1092). In a follow-up study, Goldman et al. (2012) found qualitative differences between better and poorer learners’ evaluations of text reliability. In particular, college students who learned more from their reading of multiple texts were slightly more likely to cite author credentials (e.g., ‘it’s by a

professor, so it's probably accurate' or 'this guy's a crackpot'), address information quality (e.g., "that's a lot of information about plants" or "good analogy") and use the scientific soundness of the information (e.g., no evidence, proof here – pretty far-fetched") to judge texts as trustworthy or not (p. 368).

Work by Braasch and colleagues (2013) returned similar findings. They designed a classroom intervention that leveraged contrasting case analyses and ratings of website reliability that was found to support the development of evaluation skills. After generating their own ideas about effective evaluation strategies online, treatment students (average age, 17.94) were asked to read transcripts of the online reading strategies of students, ostensibly from another school, in order to distinguish what good readers do differently from struggling readers (p. 186). In truth, the contrasting cases were constructed by the research team to include stark differences. One case demonstrated better strategies, the other demonstrated weaker ones. Students read the cases independently and then discussed their evaluations with partners before a whole-class discussion, during which students decided on strategies they could use themselves when reading to learn from multiple texts. The next day, students were given 30 minutes to read a set of six pre-selected texts for the purpose of answering the inquiry question: "Explain the causes of the typical weather patterns in the Pacific Ocean and the processes that make El Niño change these weather patterns." (p.187). They took no notes. They then had 20 minutes to write an essay in response to the inquiry prompt. After writing, students were asked to rate the six texts on a scale of usefulness. In their essays, treatment participants included more core scientific concepts from more reliable texts than control participants, who received regular classroom instruction (p. 190).

Together, findings from studies by Braasch et al. (2013), Goldman et al., 2012, and Wiley et al. (2009) indicate that critical evaluation skills can and should be taught because they seem to

support content learning. Moreover, the instructional methods used by Braasch and colleagues (2013) and Wiley and colleagues (2009) have been proven effective. What is not clear from this pair of studies is the extent to which the interventions resulted in participants constructing more *integrated* mental models of understanding. Participants who received the interventions were able to include more content from more reliable information sources in their written work, but it is not clear whether these students' essays included evidence of corroboration across multiple texts, comparisons to background knowledge, juxtaposition of contrasting facts gathered across a range of multimodal texts, or linguistic markers suggestive of integrative cognitive processing—markers of advanced multiple text integration that we might expect in the written essays of more expert multiple text integrators. If the content—the what—of an intervention were extended to include more of the expert integration strategies outlined above, would students become better multiple text integrators?

Given the fundamental importance of multiple Internet text integration skills in the digital age (e.g. DeSchryver, 2012; Rouet, 2006) I assert that it is essential to leverage knowledge of expert and novice multiple text integration processes to both devise and test instructional methods that could enable students to become more expert multiple text integrators. Additionally, methods that support younger students are critically needed.

$[(PST)^2 + (iC^3)]$

Though surely not a complete list, current theory and empirical evidence, suggests that construction of an integrated mental model of understanding from multiple, multi-modal Internet texts requires (at least) the ability to: (a) construct meaning from single texts (Kintsch, 1998), (b) find, gather, evaluate, compare and connect information from multiple texts, documents and sources (Braasch et al., 2013; Goldman et al., 2012; Leu et al., 2004; Rouet, 2006; Wiley et al.,

2009; Zhang & Duke, 2011), (c) make meaning from multi-modal elements (Jewitt, 2008; Kress & vanLeeuwen, 2010), (d) flexibly and strategically apply a range of online and offline reading comprehension skills (Afflerbach & Cho, 2009; Coiro & Dobler, 2007; Shanahan, 2009), (e) leverage background knowledge to navigate across hypertexts, and (f) use search engines, and especially cues in the SERP to determine relevance of texts and make forward inferences (Afflerbach & Cho, 2009; Coiro & Dobler, 2007; Henry, 2006).

In Table 1, I present the framework of ten strategies informed by this literature: [(PST)<sup>2</sup>] + (iC<sup>3</sup>). The order of presentation aligns with existing models of inquiry with multiple-texts (e.g., Perfetti et al., 1999; Rouet, 2006). Part one of the framework [PPSSTT], also [(PST)<sup>2</sup>], lists strategies thought to occur in advance of text selection; part two [iCCC], also (iC<sup>3</sup>) strategies are thought to occur most often after students choose a text for close reading. Part one strategies are (P) thinking about reading purpose, (P) forward thinking about pre-existing knowledge, (S) generating search terms, (S) skimming and then selecting sources/texts, (T) using knowledge of text type/genre to predict source/text relevance, and (T) evaluating/predicting trustworthiness. Part two focuses on the reading and integrating of information from each text with what is already understood (iC<sup>3</sup>). These strategies include: (i) identifying important information in each text, (C) making comparisons to background knowledge, (C) making connections to other texts, and finally, (C) continually updating what is understood. In the table, I include guiding questions for students to consider as they engage these processes.

### **How to Teach [(PST)<sup>2</sup> + (iC<sup>3</sup>)] to Ninth Grade Students**

Several studies have shown that, for adolescents, multiple text integration skills improve with practice (Strømsø, Bråten, & Samuelstuen, 2003) and instruction (Britt & Aglinskias, 2002; Wiley & Voss, 1999; Wiley et al., 2009). Wiley and Voss (1999) found that students produced

Table 1

*Summary of the  $[(PST)^2 + (iC^3)]$  Framework of Multiple, Multimodal Internet Text Integration Strategies*

Pre-Reading	<b>P: Purpose</b> What do we have to learn about? What do we have to create with this information? <b>P: Pre-existing knowledge</b> What do we already know about this topic?
For Finding, Previewing and Evaluating	<b>S: Search</b> What search terms should we use? <b>S: Source selection</b> Which of these sources looks most promising, and why? <b>T: Type of Text</b> What type of text is this? Does this help us understand more about the information it provides before we select it? <b>T: Trustworthiness</b> How trustworthy is this source?
MOUSE CLICK/CHOICE	
During Close Reading	<b>I: Identify important information</b> What information can we use to meet our reading purpose? <b>C: Compare to pre-existing knowledge</b> How does this information compare with what we already know? <b>C: Connect to other texts</b> How does this information connect with information that we have read in other texts? <b>C: Continually update understanding</b> What do we know now? What do we still need to understand to achieve our purpose?

the most integrated and causal essays in response to prompts that asked them to (a) form an argument and (b) when they needed to construct their argument from multiple sources presented on a website. The authors suggested that the prompt itself was a significant intervention and that the web-based presentation of sources may have played a supportive role in the synthesis

processes as well. It is not clear, however, whether argumentative topic prompts, plus a targeted intervention will lead to a more integrated understanding of topics than a comparison condition à la Wiley and Voss (1999) in which students receive the argumentative topic prompt and Internet access but no targeted intervention. Moreover, it is important to question how such an intervention should be designed.

The most promising instructional methods for teaching online reading processes, more generally, have found their footings in the most widely supported methods for teaching reading comprehension. In an Irish school district serving disadvantaged populations of children over a two-year time period, Dwyer (2010) used a formative and design experiment (Reinking & Bradley, 2008) to test the impact of an instructional environment that sought to “scaffold the development of effective online reading and information-seeking strategies [...], within an integrated classroom curriculum, through a series of linked interventions.” (p. 74). Importantly, students in her study worked collaboratively with peers and in groups. Her instructional methods drew heavily from (a) Guthrie’s Concept Oriented Reading Instruction (CORI) (Guthrie et al., 1996; Guthrie & Wigfield, 2000) model which combines strategy instruction with conceptual knowledge instruction in science, and methods that support readers’ motivation and engagement with texts (Guthrie, McRae & Klauda, 2007; Guthrie, Wigfield & Klauda, 2012). She also borrowed methods from Palincsar & Brown’s (1984) Reciprocal Teaching framework that emphasizes four essential comprehension strategies: predicting, questioning, clarifying and summarizing, along with more general strategic comprehension monitoring. In this model, teachers use gradual release of responsibility (Duke & Pearson, 2002; Duke et al., 2011; Pearson & Gallagher, 1983) moving from direct instruction to student-led discussions of their own reading strategies that are socially supported and positioned within learners’ zones of proximal



development (Vygotsky, 1978). Within the gradual release of responsibility model, Dwyer found three instructional strategies to be particularly supportive of online strategy development: (a) brief, but explicit strategy instruction using think-aloud techniques (Kucan & Beck, 1997; Newell & Simon, 1972) (b) adaptive scaffolding that was just-in-time and responsive to students' immediate learning needs, and (c) peer-to-peer collaboration (p. 361).

In her study of 4th and 5th grade students learning to read online, Castek (2008) also found that students taught one another online reading comprehension strategies (p. 198). Although the students relied on the teachers' instructional modeling and guidance early on, teachers' roles in this study changed quickly once students' competencies grew. Regarding appropriate levels of scaffolding, Castek recommended that teachers "consider shifting quickly from teacher scaffolding to opportunities for students to scaffold one another" (p. 198). In a pre-post study of students' learning gains, Castek also found that students who received appropriately scaffolded new literacies instruction performed better than matched controls on measures of online reading comprehension (p. 176). Importantly, the treatment group made statistically significant gains on measures of content learning but did not improve on the measures of synthesis used in her study. Castek used a concept map to measure synthesis and conceptual growth, but as operationalized, the map may not have tapped into deeper levels of knowledge construction (p. 178) that might have enabled students to build a more integrated understanding across multiple texts. It might also have been the case that the task prompts used to measure online reading were not sufficiently "argumentative" so that students' reading strategies were not ever particularly geared toward synthesis of multiple texts, ideas, or perspectives. Although one task asked students to develop an opinion, younger students, like the fourth and fifth-grade participants in Castek's research might have needed more practice or instruction that was very specifically directed

toward synthesis skills development for gains to have been seen on this variable.

The Teaching Internet Comprehension to Adolescents (TICA) project (Leu & Reinking, 2005a), the goals for which have been to increase the use of Internet reading comprehension strategies to concomitantly improve (a) reading online and offline, (b) academic engagement and (c) achievement among middle-schoolers at risk of dropping out (Leu & Reinking, 2005b) has also adopted a version of Palincsar & Brown's (1984) reciprocal teaching model (Leu, et al., 2008) with promising results. As measured by specific Online Reading Comprehension Assessments (ORCA), scores on a paired-samples t-test for treatment students who received the Internet Reciprocal Teaching (IRT) intervention were significantly higher in the second year of the TICA study (Leu et al., 2008, p. 333). Consistent with Castek and Dwyer's instructional methods, IRT also prescribes teacher-led instruction, collaborative modeling of specific online reading comprehension strategies, and gradual release of responsibility until students engage in their own online inquiries (Leu et al., 2008, pp. 328-330).

The evidence presented here strongly supports an integrated, gradual release of responsibility model for online reading instruction that includes teacher modeling, responsive scaffolding, peer collaboration and opportunities for student inquiry. Although intervention studies have shown general gains in online reading comprehension skills (Castek, 2008; Dwyer, 2010; Leu et al., 2008) the particular constellation of strategies (i.e, the what) and instructional methods (i.e, the how) that might best support the development of multiple, multi-modal Internet text integration skills has yet to be determined. Given the complexities of every classroom ecology, and the complexities of multiple text Integration processes, a one-size-fits-all solution is not the goal of this work. However, there is value in articulating and testing a set of methods known to support gains in offline and online reading comprehension skills for their value as

supports for the acquisition of multiple, multimodal Internet text integration skills.

## **LINKS**

Based on an examination of promising methods for offline and online reading comprehension instruction, the LINKS intervention included seven integrated instructional elements, implemented in the following order: (a) discussion of reading prompt, reading purpose and background knowledge between students in dyads; (b) quick, direct introduction and review of  $[(PST)^2 + (iC^3)]$  strategies and questions by teacher; (c) teacher modeling of strategy use for the purpose of constructing an integrated understanding of topics from multiple texts via a series of three screencasts that gradually released responsibility to students over three intervention sessions; (d) 30 minutes of dyadic online inquiry; (e) guided teacher questioning that prompted application of  $[(PST)^2 + (iC^3)]$  strategies; (f) note taking that required students to change ink color to delineate information gathered from different information sources (not necessarily texts, since some information sources included multiple, multimodal texts – e.g., words, graphs, images); (g) writing a persuasive argument independently for 20 minutes.

I've called this intervention Learning to Integrate InterNet Knowledge Strategically (LINKS). The acronym articulates the intervention's purpose. Knowledge, in this case, stands for the schemas students build from the processes of gathering, evaluating and integrating information from multiple texts. The word LINKS is synonymous with integration, or synthesis and connotes the Internet's fundamental property—the hyperlink, often *link* for short.

## **Research Questions**

With evidence from theory and empirical research guiding both the content (i.e., the what) and the method (i.e., the how) of this intervention, this study asked the following two questions:

1. What impact, if any, does the LINKS intervention have on students' application of multiple, multi-modal Internet text integration skills during online inquiry?
2. What impact, if any, does the LINKS intervention have on trace evidence of integration processes in students' written persuasive arguments?

## CHAPTER 3

### Method

#### Design

To explore LINKS' impact on (a) application of strategic processes during reading, and (b) evidence of integration in students' persuasive arguments, I used a repeated measures design with one control group and one treatment group. Dyads were randomly assigned to treatment condition.

All participants completed five online inquiry sessions. Pretest (session 1) and posttest (session 5) followed the same format for both groups. For the treatment group, LINKS was administered during the three intermediary sessions (sessions 2, 3 and 4). I refer to these three intermediary sessions generically as "practice sessions" since, for the Control group, these three sessions were opportunities to repeatedly practice their approach to online inquiry. The treatment participants also practiced online inquiry during these sessions, but received added supports.

Table 2 outlines the design of the study for treatment and control participants. Protocols for treatment and control sessions are described, in this chapter, below.

Importantly, the study used a standard dyadic design (Kenny, Kashy & Cook, 2006), meaning that each participant was part of one dyad for the duration of the study. Non-independence was assumed during online inquiry and members of dyads were considered indistinguishable. Dyads completed all reading and online inquiry processes together, but wrote persuasive essays independently.

I also conducted a case study (Yin, 1989) of the treatment dyad whose pretest writing scores were lowest. This analysis allowed me to focus specifically on the ways the intervention influenced the participants who showed the least preparation for the inquiry and writing activities

at pretest.

Table 2

*Summary of Study Design and Activities for Treatment and Control Conditions*

Session	Treatment	Control
1:	Assent	Assent
Introduction	WRMT-III	WRMT-III
	Partner Survey	Partner Survey
2: Pretest	Topic: Soaps/Radiation	Topic: Soaps/Radiation
	Inquiry (30 minutes)	Inquiry (30 minutes)
	Writing (20 minutes)	Writing (20 minutes)
3: Practice 1	Topic: Wind/Nuclear/Math	Topic:
	<i>Introduction &amp; direct instruction of <math>[(PST^2 + iC^3)]</math></i>	Wind/Nuclear/Math
	<i>Discuss Pre-Existing Knowledge of Topic</i>	<i>Silent reading</i>
	<i>Pre-Existing knowledge on transparency layer 1</i>	<i>screencast</i>
	<i>Think aloud screencast</i>	<i>Optional note taking on</i>
	<i>Inquiry (30 minutes) [with guided questioning]</i>	<i>transparency layer 1</i>
	Note taking with color & transparency layer 2	Inquiry (30 minutes)
	Writing (20 minutes)	Note taking with color
		on transparency layer 2
		Writing (20 minutes)
4: Practice 2	<i>Review <math>[(PST^2 + iC^3)]</math></i>	Topic:
	Topic: Wind/Nuclear/Math	Wind/Nuclear/Math
	<i>Discuss Pre-Existing Knowledge of Topic</i>	<i>Silent reading</i>
	<i>Pre-Existing knowledge on transparency layer 1</i>	<i>screencast</i>
	<i>Think aloud screencast</i>	<i>Optional notetaking on</i>
	<i>Inquiry (30 minutes) [with guided questioning]</i>	<i>transparency layer 1</i>
	Note taking with color & layer transparency 2	Inquiry (30 minutes)
	Writing (20 minutes)	Note taking with color
		on transparency layer 2
		Writing (20 minutes)

Table 2 (cont'd)

5: Practice 3	<i>Review [(PST<sup>2</sup> + iC<sup>3</sup>)]</i>	Topic: Wind/Nuclear/Math
	Topic: Wind/Nuclear/Math	Silent reading screencast
	<i>Discuss Pre-Existing Knowledge of Topic</i>	<i>Silent reading screencast</i>
	<i>Pre-Existing knowledge on transparency layer 1</i>	<i>Optional notetaking on</i>
	<i>Think aloud screencast with students actively</i>	<i>transparency layer 1</i>
	<i>identifying strategy application</i>	Inquiry (30 minutes)
	<i>Discuss strategies application</i>	Note taking with color on
	Inquiry (30 minutes) <i>[with guided questioning]</i>	transparency layer 2
	Notetaking with color & layer transparency 2	Writing (20 minutes)
	Writing (20 minutes)	
6: Posttest	Topic: Soaps/Radiation	Topic: Soaps/Radiation
	Inquiry (30 minutes)	Inquiry (30 minutes)
	Writing (20 minutes)	Writing (20 minutes)

---

*Note:* WRMT-III = Woodcock Reading Mastery Test (3<sup>rd</sup> Ed.) Reading Passages Subtest (Woodcock, 2011); PST2 + iC3 = Strategic Framework that includes 10 strategies: Purpose, Pre-existing knowledge, Search, Source Selection, Type, Trustworthy, Identify Important Information, Compare, Connect and Continually Update. *Italics* used to indicate differences in group activities. Inquiry for both groups was done on the Internet and in consistent dyads.

## Participants

Results for sixteen purposefully selected ninth-grade participants matched in eight consistent, indistinguishable dyads (Kenny, Kashy & Cook, 2009) are reported in this study. Participants were recruited from two schools—one public and one independent—in a Midwestern state. In 2012-2013, the independent school had an enrollment of 540 students from grades 6 through 12, 80 of whom were in the ninth grade. The total enrollment of the public high school was 706, 188 of whom were in the ninth grade.

I chose a sampling frame of ninth-grade students for three important reasons. First, cognitive, developmental and neurological evidence suggests that adolescents experience

tremendous growth in their capacity for complex problem solving (e.g., Blakemore, & Choudry, 2006; Geidd, 2004; Gogtay et al., 2004; Kuhn, 2006) and that this special time in development might even be a sensitive period for the development of enduring higher order thinking skills (e.g., Steinberg, 2005). Given this, it is important for teachers to understand how to set students on productive learning trajectories for the development of these skills at this time in their growth. Secondly, the CCSS set ambitious standards for students in terms of integration by the end of high school. Given the changing educational context in which ninth graders find themselves at present, I saw particular value in working with students whose educational experiences will most likely be shaped by new curricular expectations for integration of texts across academic disciplines. Thirdly, as outlined in the review of literature, more research on the instructional methods and content that might support multiple text integration for younger students is needed.

All ninth-grade students in both schools were invited to participate through an IRB approved information packet sent home to families. In total, 30 students returned permission forms. Twenty-two students completed all parts of the study in consistent dyads; three completed the study as individuals after their partners opted out.

Eight dyads were purposefully selected for the current analyses so that control and treatment groups were balanced on important factors: pretest online reading scores, school, and self-reported racial/cultural identity. The gender distribution of the purposefully selected dyads, 11 girls, 5 boys, reflects the general gender disparity in the larger sample (14 girls and 8 boys completed the study). The control group included three girl-girl dyads and one boy-boy dyad. The treatment group included one boy-girl dyad, two girl-girl dyads, and one boy-boy dyad. Table 3 summarizes participation by school.

The average age of participants in the eight dyads at the start of the study was 14 years,



eight months (or 14.67). On a self-report survey, 11 students self-identified as white/Caucasian, three as Black/African American, one as South-Asian and one as Persian/Middle Eastern. All minority students attended the independent school. Table 4 provides a summary of descriptive participant data.

Self-report data for the eight dyads suggests participants were generally familiar with the Internet. All participants reported Internet access at home, and at school. At school, 14 (87.5%) participants reported using Google searches to find information about topics, and visiting websites in school for specific purposes as directed by a teacher. Eleven (68.75%) reported using library resources such as online databases to find information for projects. When asked to report what they did most frequently at school on the Internet, results were mixed. Five participants reported creating multimedia presentations most frequently (31.25%), four reported searching for information (25%), three reported visiting teacher recommended websites (18.75%), and one reported online discussion (6.25%) was the most frequent online activity in which she engaged at school. Three participants did not answer this section (18.75%) of the survey.

Table 3

*Summary of Participation by School*

	Boys	Girls	Treatment Dyads	Control Dyads
School 1 (n=4)	2	2	1	1
School 2 (n=12)	3	9	3	3
Total	5	11	4	4

*Note:* School 1 = public high school; School 2 = independent school

Table 4

*Summary of Descriptive Data for Control and Treatment Dyads*

	Dyad	WRMT	DYAD MEAN WRMT	Gender	School
Control	1	89	98	F	2
	1	107		F	2
	2	107	118	F	2
	2	129		F	2
	3	125	129.5	M	2
	3	134		M	2
	4	117	121	F	1
	4	125		F	1
Treatment	5	110	117.5	F	2
	5	125		M	2
	6	101	109	F	2
	6	117		F	2
	7	117	123	F	2
	7	129		F	2
	8	104	119	M	1
	8	134		M	1

*Note:* School 1 = public high school; School 2 = independent school, WRMT = Woodcock Reading Mastery Test Passages Comprehension Subtest.

The picture was slightly different for home Internet use. Six of 16 participants (37.5 %) reported engaging in social networking such as Facebook or Twitter most often at home. Three (18.75%) reported doing homework most often, two (12.5%) participants reported gaming as their most frequent activity, two (12.5%) reported watching videos online and one (6.25%) reported doing email most often. Interestingly, six participants (37.5%) reported that doing

research for school projects was either their second or third most frequent online activity at home. Among second and third choices, social networking was cited by five participants (31.25%) and watching videos was cited by six participants (37.5%).

### **Research Context**

The independent school curriculum emphasizes college preparation but is not required to deliver state curriculum or administer state standardized tests. Admission is selective but based on several factors including the student's academic and extra-curricular interests, teacher recommendations and test scores. The student body is purposefully selected to reflect the racial, economic, religious and social diversity of the surrounding community, which, according to United States Census Bureau statistics (2012) indicate 73% of the population is white, 7.7% African American, 14.4% Asian and 4.1% Hispanic or Latino/a. Tuition payments at the independent school are prorated to family income in order to ensure access for all accepted applicants. The median family income from 2008-2012 in this city was \$53, 814 (United States Census Bureau, 2012) which is above the statewide median of \$48, 471. In this city, 21.9% of persons live below the poverty level versus 16.3% statewide. Free-reduced lunch data was not available for this school.

The public school, on the other hand, is expected to follow the State curriculum and administer all required State standardized tests to students. Ninety-four percent of the students who attend this school are white/Caucasian, which is consistent with the city's racial composition (United States Census Bureau, 2012) of 95% white, 1.1% African American, 3.1% Hispanic/Lationo/a and .7% Asian. The median household income in this city is \$49,167, slightly above the statewide median. Overall, 22% of students at this school qualified for free/reduced lunch (Center for Educational Performance and Information, 2013).

Given the sensitive nature of socio-economic data, I did not ask families or students to report their income levels and decided, instead, to use state and national data sets to contextualize the socio-economic composition of the sample. From these data, we can infer that the sample is taken in contexts where the median income is higher than the statewide median.

Both schools were equipped with high-speed Internet via wifi in all classrooms. Both schools provided laptops on carts for teachers to use in classrooms. Desktop computers were available for student use in media information centers and computer labs in both schools.

In both schools, students were excused from classes to participate in the research study. Efforts were made to ensure students did not miss the same class more than once for the duration of the study. This also controlled for time of day as a factor in the study. When scheduling permitted it, two dyads of participants attended at the same time but were positioned in the room so that their conversations did not interfere with one another. This choice was taken to reduce anxiety among participants and to provide a research context that more closely approximated a classroom, but allowed me to monitor activities closely. I hypothesized that it would also be logistically helpful since I could collect data with more students during the same blocks of time. Student absences, illnesses, tests, and other responsibilities did, however, impact the data-collection schedule, forcing some dyads to participate at times when no other dyads were available.

The study was conducted in the media center conference room at the public high school. The independent school provided an empty science lab space for the first few weeks of the study and an open conference room for the last two sessions of the study. Participants used one shared computer for all screencast and online reading activities. They used separate computers for writing.

## **Assignment to Dyad**

Dyads were assigned on the basis of two factors: (a) similarity of scores on the Woodcock Reading Mastery Passages Comprehension Subtest (version III) (Woodcock, 2011) (WRMT), and (b) students' given preferences of partner. I matched partners by WRMT score first and then adjusted based on surveyed preferences so that in each dyad, at least one of the members listed the other as one of four preferred partners. This scheme worked for all but one group. Dyad 3 were the highest scoring male readers in the sample. Their preferred partners were better matched on WRMT scores with other students so I placed these boys together, even though they had not requested to work together.

The rationale for this approach was based on evidence that several factors could influence the success of students' collaborative reading activities, including (but not limited to) (a) background knowledge of the topics (Coiro, 2011a; Kintsch, 1998); (b) interest in and motivation to read about the topics (Guthrie & Wigfield, 2000; Guthrie, Wigfield & Klauda, 2012); (c) offline reading comprehension skills as measured by standardized tests (Coiro, 2011a); (d) online reading proficiency (Coiro, 2011a); and (d) the degree to which students trust or like their partner (Dirks, 1999; Kiili, Laurinen, Marttunen, & Leu, 2012). Of these factors, I prioritized offline reading scores because of evidence that suggests it is a statistically significant predictor of online reading (Coiro, 2011a). Importantly, dyadic strategic reading has been found to support comprehension, generally (Vaughn, et al., 2011).

## **Sequence and Pacing of Research Sessions**

As outlined in Table 2, all participants met with me six times. During the first session, participants were introduced to the study, were reminded of their rights as participants and asked for their voluntary assent. They also completed the WRMT (Woodcock, 2011) and indicated

their preferences of reading partners. Session 2 was the pretest, sessions 3, 4 and 5 were practice/intervention sessions. Session 6 was the posttest session.

Sessions generally occurred at one-week intervals. However, the independent school exam schedule, student illnesses and scheduling conflicts through the late fall in both schools forced a three-week delay between the second and third practice sessions for most of the participants. Only Dyad 4 completed the study in seven weeks. Six others completed the six sessions over 10 weeks. Dyad 5 needed an extra week because of scheduling conflicts. The intervention and duration of the time between pre- and posttesting were therefore influenced by the very real context of the school year, with disruptions and breaks due to individual and institutional factors.

## **Instruments**

**Reading passages subtest of the Woodcock Reading Mastery Test.** Given that offline reading comprehension skills predict online reading comprehension skills (Coiro, 2011a), students completed Form A of the Passages Comprehension Subtest of the WRMT-III (Woodcock, 2011). This instrument, developed with a nationally representative sample of 5,000 participants aged 4 to 79 between 2009 and 2010, measures reading comprehension. It presents a series of cloze-type passages that test “an examinee’s ability to study a sentence or short passage and exercise a variety of comprehension and vocabulary skills in identifying a missing word” (p.4). Passages are designed so that readers must understand the whole passage in order to provide the correct response. In terms of the instrument's construct validity, Kintsch's Construction-Integration model (1998) would predict that the cueing of word meaning and contextual meaning in this type of cloze task engages essential reading comprehension processes. Minimum raw score on the test is 0. Maximum raw score on the test is 38. The mean raw score

for students in the first semester of 9th grade is 27, which corresponds to a standardized mean of 100.4 (SD=16.7) (Woodcock, 2011, p. 97). Reliability for Form A of the Passages subtest, calculated using the split-half method, is high ( $r = .90$ ) for 9th graders (Woodcock, 2011, p. 107).

**Recorded audio, video, and navigational clickstream data.** To investigate the intervention's impact on students' use of  $[(PST)^2 + (iC^3)]$  strategies (RQ1), I recorded their online inquiry activities using Morae Recorder screencapture software (Techsmith, 2012) on laptop computers. Audio, video, and navigational clickstream data were recorded. Inquiry recordings, each approximately 30 minutes in length, were then imported to Morae Manager (Techsmith, 2012) where they were transcribed and coded for evidence of strategy use.

**Recorded notes and background knowledge.** Students were given transparency film on which to take notes, using different colored pens, during their reading sessions. Treatment participants were told to write their background knowledge on one transparency film and to layer their second transparency over top so that they could see their background knowledge as they took notes. This also allowed me to know what students knew in advance as I evaluated their essays. Control participants used the first transparency sheet to record ideas during silent reading of "starter texts" through their screencast viewing time, but were not explicitly instructed to record their background knowledge. Since it was hypothesized that awareness of background knowledge would promote integration of multiple texts in the treatment condition (e.g., Anderson & Pearson, 1984; Kintsch, 1998; McNamara & Shapiro, 2005) I did not ask control participants to rate their background knowledge or record it in any way. Control participants also wrote on transparency film but were not explicitly asked to write background knowledge on their first sheet.

**Written persuasive arguments.** To explore the impact of LINKS on students' construction of an integrated mental model of understanding, trace indicators of integrative processing (TII) in students' written persuasive arguments were examined. After reading with a partner for 30 minutes, each participant wrote independently for 20 minutes. Participants wrote in Google documents on individual laptop computers. Persuasive arguments were saved in a folder and labeled with ID codes. Essays were evaluated for trace indicators of integration (TII) and overall quality, as described below. The choice to evaluate essays for TII and overall quality was taken because of questions that emerged through the research process about whether, or to what extent overall quality of persuasive argument and integration might be confounded.

***ACT persuasive argument scoring guidelines.*** Persuasive arguments, scored blindly to session and condition, were evaluated for evidence of genre-specific quality using a persuasive essay rubric authored by ACT (2006) and made available to teachers on the State Department of Education website (see Appendix A). Analyses of ACT rubric scores, although not specifically aligned with the research question on integration, were used to explore any movement in general genre-specific quality in the groups. It was not clear to me what the impact, if any, of the LINKS intervention might be, so I included an analysis of the students' essays as a secondary data source and point of comparison for the TII Index. I chose the ACT rubric because the ACT is included as part of the state's standardized testing program for 11<sup>th</sup> and 12<sup>th</sup> grade students and could therefore influence instruction, particularly in public schools. Although I could not confirm this, I hypothesized that if participants had been taught how to write persuasive essays in school, that the criteria outlined in this particular rubric may have influenced instruction. Criteria of evaluation on this rubric include position and understanding of task, complexity, focus and development of ideas, organization and language. The minimum score on this rubric is 0, the



maximum score is 6. I tested the reliability of the scores through Cohen's kappa analyses of interrater agreement. These are reported in the Results chapter.

***Trace indicators of integration rubric.*** I developed a rubric to score students' persuasive essays for trace indicators of integration (TII). The rubric included ten criteria, all scored on a three point scale (0, 1, 2). The minimum score on the rubric was 0, the maximum 20. The rubric is provided in Appendix B.

The TII index is an extension of the Online Reading Comprehension Assessment (ORCA-Open) developed by Leu, Coiro, Kulikowich, Sedransk, Everett-Cacopardo, McVerry et al. (2012). In addition to measuring *Reading to Locate*, *Evaluate* and *Communicate* information online, the ORCA-Open measures evidence of intertextuality, and integration of details from two websites. It gives credit for the inclusion of a claim, with evidence, using two relevant details in a written product (p. 14). The TII index that I developed aligns with the LINKS intervention in that it gives credit for use of background knowledge and multimodal texts, but also extends the conceptualization of what integration might involve for ninth-grade students. The TII index includes use of multimodal texts as a criterion (e.g., video, images, graphics, text) because this would indicate integration of ideas from multiple semiotic systems (Jewitt, 2008; Kress & van Leeuwen, 2001). The rubric gives points for evidence of connections to pre-existing content knowledge because domain knowledge is a critical predictor of comprehension (Anderson & Pearson, 1984; Kintsch, 1998; McNamara & Shapiro, 2005). Another criterion is the use of contradictory facts from one or more informational sources. This is included as a proxy for the processes of gathering relevant information that enables the construction of an argument, which is the stated purpose for each of the inquiry tasks. The rubric also includes points for use of linguistic markers that demonstrate integration of ideas from multiple texts (e.g., seriation,

transitional phrases that connect or juxtapose ideas) as a way of including the role that writing may play in integration (e.g., Langer, 1986a). Côté and Goldman (1999, 2004), Côté, Goldman and Saul (1998), and Goldman and Wiley (2004) have modeled this method of discourse analysis as a way to make inferences about integrative processing. Importantly, these TII criteria were bundled as a way to access several processes that may be constituents of multiple Internet text integration as a whole, and that may occur during inquiry. I hypothesized that integration of ideas might happen as students engaged in this range of activities during online inquiry and as they constructed their written arguments. The rubric therefore reflects his hypothesis.

Reliability of TII scores was established through a process of interrater agreement, represented by Cohen's kappa, and described below. Importantly, the process of evaluating integration accurately on the TII index depended on me—the primary evaluator —having carefully watched and transcribed students' online inquiry processes, making note of each website they visited and identifying the information source for each written note the students made on their transparency films. In this way, I was able to identify the provenance of information that appeared in each persuasive argument, composed immediately after the inquiry session. Clues about background knowledge, for instance, were captured on treatment participants' transparencies, but also from students' conversations during inquiry. Though arduous, the process of evaluating trace evidence of integration followed from each of the texts each student saw, read, and annotated in their notes, and from the conversations they had with their partner.

### **Conditions of Treatment**

**Treatment condition.** As noted above, treatment condition dyads received the LINKS intervention which included: (a) dyadic discussion of reading prompt, reading purpose and the

recording of background knowledge on the topic before any other activity; (b) quick, direct introduction to or review of  $[(PST)^2 + (iC^3)]$  strategies; (c) teacher modeling of strategy use for the purpose of constructing an integrated understanding of topics from multiple texts via screencasts that gradually release responsibility to students over three intervention sessions; (d) 30 minutes of dyadic online inquiry; (e) guided teacher questioning that prompted application of  $[(PST)^2 + (iC^3)]$  strategies; (f) note taking that required students to change ink color to delineate information gathered from different information sources; (g) writing a persuasive argument for 20 minutes. Teacher questions during online inquiry sessions were tailored to students' reading activities but always focused on the use and development of  $[(PST)^2 + (iC^3)]$  strategies.

**Control condition.** In contrast, the control condition included no strategies instruction, modeled think aloud during screencasts or guided questioning based on the framework of  $[(PST)^2 + (iC^3)]$  strategies. Instead, they (a) received the topic prompt without instruction to discuss it, (b) watched a screencast that presented texts but included no think-aloud modelling of strategic processes, (c) read together on the Internet for 30 minutes, receiving only simple teacher check-ins to see how things were going, or to address technical issues (d) took notes on transparency film, changing color by information source, and (e) wrote independently for 20 minutes.

### **Screencasts**

Both control and treatment conditions watched screencasts. The decision to provide think aloud modeling (Kucan & Beck, 1997; Kymes, 2005; Newell & Simon, 1972) via screencast was taken to ensure consistent delivery for all participants but also methodological transparency. All screencasts introduced content connected to the reading prompt. All screencasts included one website and one video. Scripts for all treatment and control screencasts used the same general

format (see Appendix C for a script example) with variations related to topic. Variation in total screencast length did occur at each session for different topics because the videos, included as examples of multimodal texts, were of different length for each topic. Pacing, too, was variable.

Control participants were told to read along silently as I browsed through “starter” texts for them. Treatment participants listened to think-aloud modeling of [(PST)<sup>2</sup> + (iC<sup>3</sup>)] strategies as I navigated the same texts as in the control screencasts, in the same sequence.

Both groups had access to the same content for the same amount of time. Treatment screencasts included think aloud notetaking, which added extra time to their screencasts.

Times also varied by session since the focus of the treatment screencast differed according to the gradual release model. Content, however, did not. Regardless of session, the same texts, presented in the same order, were used for all participants.

The total amount of time spent interacting with screencasts, by dyad, is summarized in Table 5. Regression analysis showed that screencast exposure time (in seconds) was not a statistically significant predictor of total strategic processing episodes at posttest  $F(1,7) = 0.024$ ,  $p = .882$ ,  $R^2 = .004$ , nor of the sum of integration indicators observed in essays 2, 3 and 4  $F(1, 15) = 0.551$ ,  $p = .470$ ,  $R^2 = .038$ . Total screencast time was not a statistically significant predictor of posttest integration scores either  $F(1, 15) = .141$ ,  $p = .713$ ,  $R^2 = .010$ .

For all groups in the study, topics of inquiry were counterbalanced to control for reading order as an influence on the outcome. Screencasts on each topic were therefore revised to reflect the point in the study at which students engaged with each topic. Although the content for screencasts on each topic was identical, the think aloud modeling (treatment) or amount of silent reading time (control group) for each screencast differed according to the stage of the intervention plan.

For the treatment condition, screencast 1 introduced all 10 reading strategies. Screencast 2 focused on the second part of the  $[(PST)^2 + (iC)^3]$  framework—identifying important information, comparing, connecting and continually updating. Screencast 3 modeled the reading process, but with less explicit naming of the  $[(PST)^2 + (iC)^3]$  strategies. During the third training session, treatment group participants were asked to identify the strategies that they observed in the screencast as a “review”. They were asked to recall the strategies observed in the screencast with their partners before starting their own online inquiry process. In this way, responsibility for thinking about strategies was released to the students. The list of screencast URLs can be found in Appendix D

Table 5

*Summary of Screencast Times*

Dyad	Total Screencast Time (in minutes)	Total Screencast Time (in seconds)
1	20:08	1208.00
2	20:08	1208.00
3	18:06	1208.00
4	18:21	1208.00
5	31:59	1086.00
6	34:19	1086.00
7	31:59	1338.00
8	34:06	1338.00

**Guided Questioning**

**Treatment condition.** As participants engaged in their online inquiry processes, I prompted treatment dyads to focus their attentions on the strategies outlined in the  $[(PST)^2 + (iC)^3]$  framework. Questions, though implemented at variable times because they had to be responsive

to participants' processes and needs were generally the questions outlined in Table 1. I tried to provide the most questions during practice sessions 1 and 2 with a little less during session 3, although given the variability based on students' needs this was not held constant. For the treatment condition, questions sometimes lead to follow-up questions. Sometimes students asked follow up questions or asked for help at which point, I scaffolded their integrative thinking as appropriate. I provide examples of two exchanges below.

*Example 1: Dyad 5, Amita and Jacob [Treatment Group]*

**Note: I am M for all conversations.**

M: How's it going guys?

Jacob: Okay. So, we're looking at this one. We're looking at it from a critic's point and apparently industrial wind turbine farms, this guy is like reasoning that they actually are adding to the problem.

M: Okay.

Jacob: So, I. We kind of wanted to see what this video was about.

M: Ya.

Jacob: To more of an idea from the critic's side of it.

M: Hmmm. Do you think that it's reliable information?

Jacob: That's what we're trying to figure out.

*Example 2: Dyad 7, Jane and Sharon [Treatment Condition]*

[Having just watched a video about Chernobyl, the girls summarize their understanding]

Jane: Okay, so it released radioactive

Sharon: High radioactive levels, like radiation levels.

Jane: Ya.

M: Do you see connections between what you've just listened to there and what you've read or seen anywhere else?

Jane: We haven't really seen much about the disasters, so.

Sharon: Ya.

M: Good. Well, what you've read is an important part of developing your argument. So, it's good that you have that new perspective on the issue.

Sharon: I didn't know that Chernobyl was real. I thought it was just a movie thing.

M: You thought it was just what, Sharon?

Sharon: just that movies that they had come out

M: Oh, ya. No. It happened when I was a kid. And you brought up though, in Japan, after the tsunami, you brought that point up yourself. I overheard you say it. They're similar.

Well, different cause, but similar circumstances for the poor people who were affected by it.

[while Michelle scaffolds connections across texts, Jane uses back button to return to Google search]

**Control condition.** To control for instructor presence, I did interact with the control participants periodically so that they knew I was in the room by providing friendly check-ins to just “see how things were going”, inform them of the amount of time left to read and to just ask what they were reading if they hadn't said much for a while. I did not, however, scaffold connections, or provide guiding questions that focused on strategies use. I share two representative examples of exchanges from control dyads to demonstrate the nature of these conversations.

*Example 1: Dyad 2, Alyssa and Meredith [Control Group]*

M: How's it going?

Alyssa & Meredith (together): Good.

M: Well, you've got a lot of information.

Alyssa: Oh my god.

[they navigate back to Google, Alyssa types new search term: Three Mile Island]

*Example 2: Dyad 4, Jennifer and Lori [Control Group]*

M: [seeing the computer screen] That's really small print. I think that you can zoom in. I think that you can go like this [leaning in to show them how to zoom on the track pad] to zoom it.

Jennifer: [tries it] Oh, I see.

Lori: Thank you.

M: I think we've figured out the small print problem!

Jennifer: Ya [nodding in a way that suggests she was bothered by the small print on the screen]

## **Topic Prompts**

Topic Prompts are provided in Appendix E. The reading topics (See Appendix C for the list) were taken directly from the state science curriculum for high school (Michigan Department of Education, 2006). The prompts were written to encourage an argumentative stance because this type of prompt has been found to support multiple text integration (Perfetti et al., 1999; Wiley & Voss, 1999). Further, argumentative writing in science and technical subjects is an expectation laid out by the Common Core for grades nine and ten. To balance interest and background knowledge for different domains of science, I chose one topic prompt each from the curricula for biology, earth sciences, chemistry and physics.



## Data Analyses

Four data sources were used to inform the research questions: (a) integrated clickstream and picture-in-picture video recordings of students reading in dyads, and thinking aloud while on the Internet; (b) persuasive essays written immediately after reading; (c) time stamp sequences, collected in Morae (Techsmith, 2012) at the onset of each code marker; and (d) WRMT-III scores on the passages completion sub-test.

Video and essay data were transformed through qualitative methods of coding so that they could be analyzed quantitatively. Creswell (2009) describes this approach as “integrated mixed methods”(p. 208) because qualitative techniques are leveraged to conduct quantitative analyses. Coding schemes, their development and protocols for application are reviewed below.

**Data analyses for research question 1.** Given the focus of the first research question, which was to understand the impact, if any, of the intervention on students’ ability to apply multiple, multi-modal Internet text integration skills during online inquiry, inquiry process data were analyzed for evidence of strategy use, and particularly for those strategies that were explicitly taught to students in the treatment group. Strategic episodes (defined below) were analyzed for frequency, relative frequency, relative duration and mean gap. Each procedure is described in this section.

**Coding manual.** The final list of codes used to categorize strategic episodes is provided in Appendix F. Appendix G includes sample excerpts that demonstrate how codes were applied. The list of codes was developed in phases. The initial set of *a priori* top-level categories aligned with the [(PST)<sup>2</sup> + (iC<sup>3</sup>)] framework. Subcategorical codes, or precisely which “activities and verbal interactions” would align with each top-level category were first informed by the protocols of Coiro, Castek & Guzniczak (2011) and Afflerbach & Cho (2009), but revised

through a process of open coding (Bazeley, 2013; Corbin, 2009) that ensured the coding scheme reflected participants' online reading and inquiry processes. This methodological choice was informed by Corbin (2009) who suggests, "Analysis should be relaxed, flexible and driven by insights gained through interaction with data rather than being structured and based on procedures" (p. 41).

Due to the constraints imposed by Morae in terms of how strategic episodes could be labeled, I had to give letter names to each strategy type. Table 6 summarizes the correspondence between the strategies in the coding manual and the way they were labeled in Morae during analyses. I have used these letters to identify each strategy in Tables and Figures and therefore summarize them here to make their meanings transparent for readers.

Table 6

*Summary of Letter Label and Strategy Correspondence*

Letter Label	Strategy
A	Purpose
B	Pre-existing Knowledge
C	Search
D	Source Selection
E	Type
F	Trustworthy
G	Identify Important Information
H	Compare to Pre-existing Knowledge
I	Connect to Other Texts
J	Continually Update
K	Student Question
M	Construct and Understanding Within a Single Text
N	Students Take Notes

Table 6 (cont'd)

P	Researcher Check-in [Control]
Q	Researcher Question [Treatment]
S	Scaffolding [Treatment]
Y	Trustworthiness During Reading

---

***Unit of analysis.*** Codes were assigned to *strategic episodes*, defined as actions, decisions, exchanges and/or explanations that appeared connected to the same strategic online reading process. Given that video, audio, and clickstream data were simultaneously analyzed for evidence of strategic processing, the decision to assign a code to a strategic episode could be based on evidence from one, two or all three of these modalities. A new strategic episode was assumed to begin when evidence for a new strategic process became evident. This choice follows from the work of Kiili (2013) who analyzed “episodes” (p. 252) in her study of collaborative online readers. She defined an episode as “a thematic entity consisting of successive activities and verbal interactions” that served one of five reading practices related to her research questions (p. 252). Kiili identified “locating information episodes”, “evaluating episodes,” “synthesizing episodes,” and “monitoring and regulating episodes”. I, on the other hand, coded for episodes consistent with the strategic [(PST)2 + (iC3)] framework. In sum, 3006 episodes were identified in the set of 40 videos recorded by these eight dyads.

This representative excerpt taken from dyad 3’s second practice session shows how strategic episodes were bounded. The ***top-level category code*** is provided in bold italics, the *duration of the episode from the timestamp* given in italics, Alan and Christopher (pseudonyms) are the students, [square brackets include contextual information, actions or interpretation of the spoken data], dialogue and search terms are written in plain text. Notice that episodes are

bounded by shifts in strategic activity. In this excerpt, Alan and Christopher are reading from a text, but Christopher clicks away from it and initiates the process of finding a new, more promising or relevant text to read. Christopher controls the keyboard and cursor for the duration of this episode so the actions are largely his alone. Notice though, that Alan suggests a search strategy at 9:56. In response, Christopher essentially paraphrases Alan's suggestion.

Paraphrasing to generate search terms could be considered a separate strategic activity and indeed, other published coding schemes have identified paraphrasing as a discrete unit of analysis (Coiro, Castek & Guzniczak, 2011, p. 369). In view of the larger strategic purpose on which this exchange centered, it was coded as a single strategic episode of *Search* but the paraphrasing was noted as a sub-category of search so that more specific comparisons of strategic activity within the meta-level categories can be conducted in future.

### ***Identify important information***

9:48.0 - 9:49.6

Alan: [reading from text as Christopher scrolls down the page] Calculating [suggesting this could be important information]

Christopher: [responds] That's just how to [and clicks away from site - he does not complete his sentence, but his response to Alan suggests he does not think this information is important]

### ***Source Selection***

9:49.6 - 9:56.0

[Back to Google]

Christopher: [controlling the cursor; scrolls through previous results, looking for a new text]

### ***Search***

9:56.0 - 10:17.8

Alan: [suggests search terms] I think we should find out how much energy it takes to produce a wind turbine.

Christopher: [types search term into search bar, dictating aloud as he types] Production of wind turbines in terms of energy? [Paraphrase]

Christopher: [clicks search]

### ***Source Selection***

*10:17.8 - 10:29.7*

Christopher: [controlling the cursor - skims and scrolls through search results]

Christopher: [reads aloud] Wind energy basics

Christopher: [continues to scroll]

### ***Search***

*10:29.7 - 10:57.4*

Christopher: [clears address bar]

Christopher:[types new search term into address bar directly] fossil fuel energy output [and then justifies choice aloud] So, we'll have something to compare it to.

Christopher:[clicks enter]

### ***Source Selection***

*10:57.4 - 11:05.3*

Christopher: [quickly skims and scrutinizes search results; scrolls down]

Christopher: [clicks on a text : [www.globe.net/articles/2012/april/3/renewable-energy-output-must-more-than-quadruple-to-replace-fossil-fuel-study](http://www.globe.net/articles/2012/april/3/renewable-energy-output-must-more-than-quadruple-to-replace-fossil-fuel-study)]

Although transcriptions indicated turn taking between interlocutors, strategic episodes were coded and recorded without division by interlocutor. The purpose of this study was not to

describe the qualitative nature of the dyadic interventions. Rather, I leveraged the theoretical assumption that students would influence one another and that generally, the negotiation and collaborative construction of understanding through online inquiry would contribute to the development of an integrated model of understanding for both students as has been shown by others (Kiili, et al., 2012 ; Coiro, et al., 2011). For the purpose of statistical analyses then, the independent “unit” is the dyad.

***Interrater agreement.*** To test the validity of the codes and the reliability of their application to the data, coding progressed through two phases of constant comparison (Glaser & Strauss, 1967; Miles & Huberman, 1994; Denscombe, 2003) and interrater agreement.

*Phase 1.* A set of 6 purposefully selected videos, three treatment, and three control group, were initially transcribed and coded to develop consistent coding methods. The first iteration of the coding manual was reviewed with an expert colleague. These discussions focused on the structure, meaning, and consistent application of the codes. The expert colleague coded 40 randomly selected excerpts. All coding differences were resolved through discussion, review of the original video data, and careful review of definitions. The refined codes and nuanced interpretations discussed during this session informed all subsequent coding of video data. Although we negotiated agreement on a random sample of codes, this phase of interrater agreement was designed to identify and resolve problems at an early stage (Bazeley, 2013) so that subsequent analyses would be more reliable. Revisions to the coding scheme based on these discussions were applied to the first six videos and to the remaining 34.

*Phase 2.* Once all video process data had been coded, the same expert colleague coded a random sample of 264 strategic episodes. Interrater agreement was very high (Landis & Koch,

1977, p.165)  $k=.874$ ,  $p<.001$ . All differences were resolved by viewing and discussing the original video evidence.

Finally, all process codes were updated and checked a third and final time to ensure consistency.

***Frequency counts.*** With all strategic episodes coded and checked, frequencies of codes were analyzed in SPSS v. 21 (IBM, 2013). This enabled comparisons of macro-level differences in the number of strategic episodes counted during each online inquiry session for treatment and control dyads.

***Relative frequencies.*** For each code during each online inquiry session, I calculated the frequency relative to the overall total. I did this by dividing the frequency for each code at each session by the total number of strategies used at that session. Differences in relative frequency within or between groups suggest different patterns of strategy application (e.g., Goldman et al., 2012).

***Relative duration.*** Using time stamps at the onset of each code in Morae, I calculated relative duration for each coding type using GSEQ (Bakeman & Quera, 2013). Duration is the measure of the total amount of time spent on each strategy. Relative duration, like relative frequency, is a proportion. For each session, I calculated the amount of time each dyad spent on each strategy, relative to the total amount of time of the session. Differences in relative duration within or between groups also suggest different patterns of strategy application.

***Mean gap.*** Mean gap is a measure of the average elapsed time between successive applications of the same code (Bakeman & Quera, 2011). It is an indicator of the quickness with which dyads applied the same strategy during the inquiry session. Longer gaps suggest a longer

time, or delay, between applications of the same strategy; shorter gaps suggest a quicker application, or shorter delay.

**Data analyses for research question 2.** To explore the impact, if any, of the LINKS instructional framework on trace evidence of integration in participants' persuasive essays, and overall essay quality, essays were coded using two rubrics. Processes are described below.

***Evaluation of persuasive arguments for overall quality.*** An undergraduate research assistant blinded all participant essays before they were evaluated for overall quality. Essays were therefore graded without knowledge of the author's identity, school, condition assignment, or the session during which the essay was written.

Two experienced English Language Arts teachers who have also conducted research on similar topics were invited to evaluate a random sample of 15 essays. Their scores were compared with mine. We found two-way agreement on 12/15 (80%) of the essays with the third rater scoring one point above or below the score given by the other two raters (i.e, adjacent agreement), however, the kappa analyses revealed very low agreement between individual raters ( $k = .022$  to  $k = .25$ ). Differences were resolved through discussion. All essays were re-coded using established protocols for consistency distilled during these discussions. Final agreement between two raters on a random sample of 15 essays was much higher,  $k = .586$ ,  $p < .001$ , a moderate level of agreement (Landis & Koch, 1977; Bakeman & Quera, 2011).

***Evaluation of persuasive arguments for evidence of integration.*** Evaluation of each essay for evidence of integration was done immediately following the coding of the think-aloud video session so that all evidence of students' reading/writing experiences were considered. Although the essays were blinded for initial coding of overall quality, this phase of essay evaluation was



not blind because (a) analyses depended on an understanding of the processes that preceded the writing and, (b) intervention was obvious in each video.

Interrater agreement for essays followed from a quantitative paradigm. I coded the essays on the TII Index rubric of 10 criteria. For each criterion, a score of 0 meant there was no evidence for that indicator; a score of 1 meant some evidence, that the essay met a minimum requirement, or that there was one example of a given indicator; a score of 2 meant there was considerable evidence, or the evidence provided exceeded minimum expectations. The total score comprised the TII Index. As I coded each essay, I provided justification for each choice using evidence from students' think alouds, clickstream data and handwritten notes. These notes were recorded directly on copies of each essay. [See Appendix H for an example of essay markup and evaluation.]

My essay evaluations were then compared with the blind evaluations of a second coder on 8 randomly selected writing samples (20% of all essays). The second coder did not watch the videos of students reading as a matter of course, nor did she review students' handwritten notes. This was reserved for when we disagreed on evaluations. She made her initial judgments about origins of ideas, number of texts used, and evidence of integration using the comments that I inserted. She did not see my scores before coding the essays herself.

Interrater agreement for the essays was within an acceptable range ( $k = .617$ ) (Landis & Koch, 1977; Bakeman & Quera, 2011). There was most agreement between us when there was no evidence or plenty of evidence for each of the criteria (codes 0 or 2) but kappa analyses revealed 5 instances when rater B gave a 2 and I gave a 1 and 6 instances when I gave a 2 and she gave a 1. This suggests a high level of adjacent agreement but it signaled a need for closer examination of differences. Differences were most pronounced on certain criteria of evaluation.

Whereas we had perfect agreement across all eight essays on evidence of background knowledge use, we were less perfectly aligned in our assessment of evidence of integration generally between texts, among texts and with background knowledge ( $k = .333$ ) On this criterion, we agreed on 2s for 5 out of 8 essays, a score of 1 for one essay but then differed in our evaluation of two essays. With only 8 examples of the application of this code, two disagreements made a significant impact on the overall evaluation of agreement. All differences were resolved through discussion and review of the evidence. Final index scores were carefully reviewed to ensure adequate evidence to support each value judgment and consistency in coding following from those discussions.

## CHAPTER 4

### Results

Analyses are reported in three sections. Section 1 explores a response to the first research question. Section 2 explores a response to the second research question. Section 3 presents a single case study that informs both questions.

To explore the LINKS intervention's impact on students' application of  $[(PST)^2 + (iC^3)]$  strategies (RQ1), I report comparisons of frequencies data, relative frequencies data, relative duration and mean gap data. Where permitted, parametric methods of analysis are applied. When data violate assumptions of normality or homogeneity of variance, results of non-parametric tests are reported.

The second section examines the LINKS intervention's impact on trace indicators of integration in students' written arguments. Where permitted, parametric methods of analysis are applied to compare within-groups and between-groups differences. When data violate assumptions of normality or homogeneity of variance, results of non-parametric tests are reported.

The third section is a case study analysis of the treatment dyad whose pretest writing scores were lowest. This analysis further informs research questions 1 and 2 by showing how the LINKS intervention impacted the reading processes and written products of the dyad who, at pretest, had the most to potentially gain from the intervention.

#### **Section 1: LINKS' Impact on Strategy Application**

**Pretest comparison of reading scores.** WRMT Passages Subtest scores for treatment and control participants were compared to determine pretest differences between groups on this validated measure of reading comprehension ability. Shapiro-Wilk tests showed that the

assumption of normality was met for treatment ( $W=0.958, p = .793$ ) and control groups ( $W = 0.925, p = 0.472$ ) on this measure. Assumptions of homogeneity of variance, as determined by Levene's test, were also met  $F(1,14) = 0.493, p = 0.494$ . An independent samples t-test was therefore justified. The null hypothesis was retained. Mean scores on the WRMT at pretest ( $n=16$ ) did not differ statistically between groups ( $t = -0.075, p = 0.942$ ). Given this finding, equivalent offline reading comprehension ability was assumed between groups. The WRMT reading score was therefore not included as a covariate in any subsequent analyses.

**Comparisons of mean frequencies of total strategy use.** To determine whether the LINKS intervention had an impact on total mean strategy application within and between groups, I first analyzed total frequencies of strategic code application. This is a macro-level test of the intervention's impact on strategic processing. Table 7 summarizes the total mean sums of processing episodes for both treatment ( $n = 4$  dyads) and control ( $n = 4$  dyads) conditions at pretest, practice session one, two, three and, at posttest. Total frequencies of strategic episodes for each dyad were defined as the sum of all  $[(PST)^2 + (iC^3)]$  codes, plus M (constructing understanding within a single text) and Y (trustworthiness during close reading) codes. M and Y codes were added because they emerged, through data analyses, as essential components of multiple text integration processes that were not adequately captured, *a priori*, in the  $[(PST)^2 + (iC^3)]$  coding categories.

Although the assumption of sphericity was met ( $\chi^2(5) = 5.566, p = .360$ ) suggesting that the variances of the differences between session levels were approximately equal, the assumption of homogeneity of variances was not met for all sessions, nor was the assumption of normality. For this reason, I compared only pretest and posttest data, using repeated measures ANOVA, to determine whether total processing events changed from pretest to posttest within groups and

between groups. The null hypothesis was retained for within groups and between groups differences. There was no statistically significant main effect of “session”  $F(1,6) = 1.048, p = .345$  within groups, meaning that frequency totals for strategic processing episodes did not differ between pretest or posttest in treatment and control conditions. Likewise, the interaction of session and condition was not statistically significant  $F(1,6) = .816, p = .401$ . For the participants in this study, it therefore seems that the type of treatment received had no statistically significant impact on the macro-level sum total of processing events applied at pretest or posttest.

Table 7

*Mean Frequencies of Total Strategic Episodes*

	Control (n=4 dyads)		Treatment (n=4 dyads)	
	Mean (SD)	95% CI	Mean (SD)	95% CI
Pretest	67.75 (18.42)	[49.70, 85.80]	59.50 (11.73)	[48.18, 70.52]
Practice 1	57.75 (13.00)	[45.01, 70.49]	59.25**(15.17)	[44.38, 74.12]
Practice 2	59.75 (20.85)	[39.32, 80.18]	73.00 (20.02)	[53.38, 92.61]
Practice 3	52.50 (5.26)*	[47.04, 57.65]*	64.75 (32.00)*	[33.39, 96.75]
Posttest	68.31 (20.03)	[48.68, 87.94]	68.50 (23.59)	[45.38, 92.09]

*Note:* Confidence intervals calculated using formula  $M \pm \left[ \frac{SD}{\sqrt{n}} * 1.96 \right]$

\*Variances not homogeneous between groups at this session

\*\*Normality assumption violated.

Descriptively, however, it should be noted that mean values for Practice Sessions 2 and 3 for the treatment group exceeded control group means by the largest margins (13.25 and 12.25 respectively), suggesting that during these sessions, when the LINKS intervention was applied, a relative increase in total strategy application occurred for the treatment dyads, compared with the control dyads. This is consistent with what one would expect, given that students were prompted to engage strategies during these practice sessions.

Given the small sample size and wide confidence intervals, a visual examination of trends in total strategy application for each dyad, compared with the mean offers some indication of total strategy application across all five sessions within and between treatment groups. Figures 1 and 2 show trend lines for each of the four dyads in the control and treatment groups across all five sessions of the study. The fifth (black) line represents the mean total. As would be expected, patterns of total processing episodes differed by dyad and by study session.

Descriptively, it seems that the treatment condition curves show fewer dramatic changes from session to session than the control group curves. This lead me to wonder if the treatment condition constrained participants' strategic processing in some way, even though mean frequency values for the groups did not differ statistically at any point in the study.

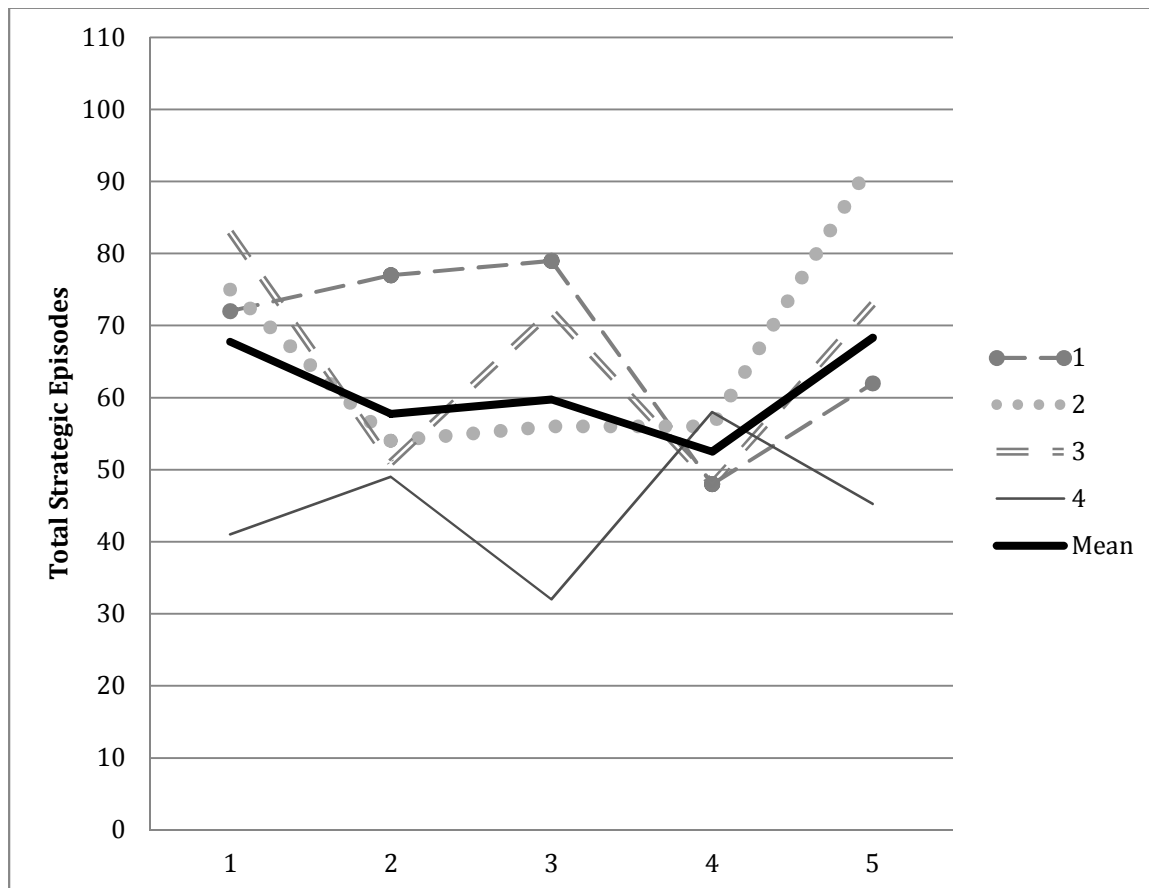


Figure 1. Total strategic processing episodes over five sessions for control dyads (n = 4).

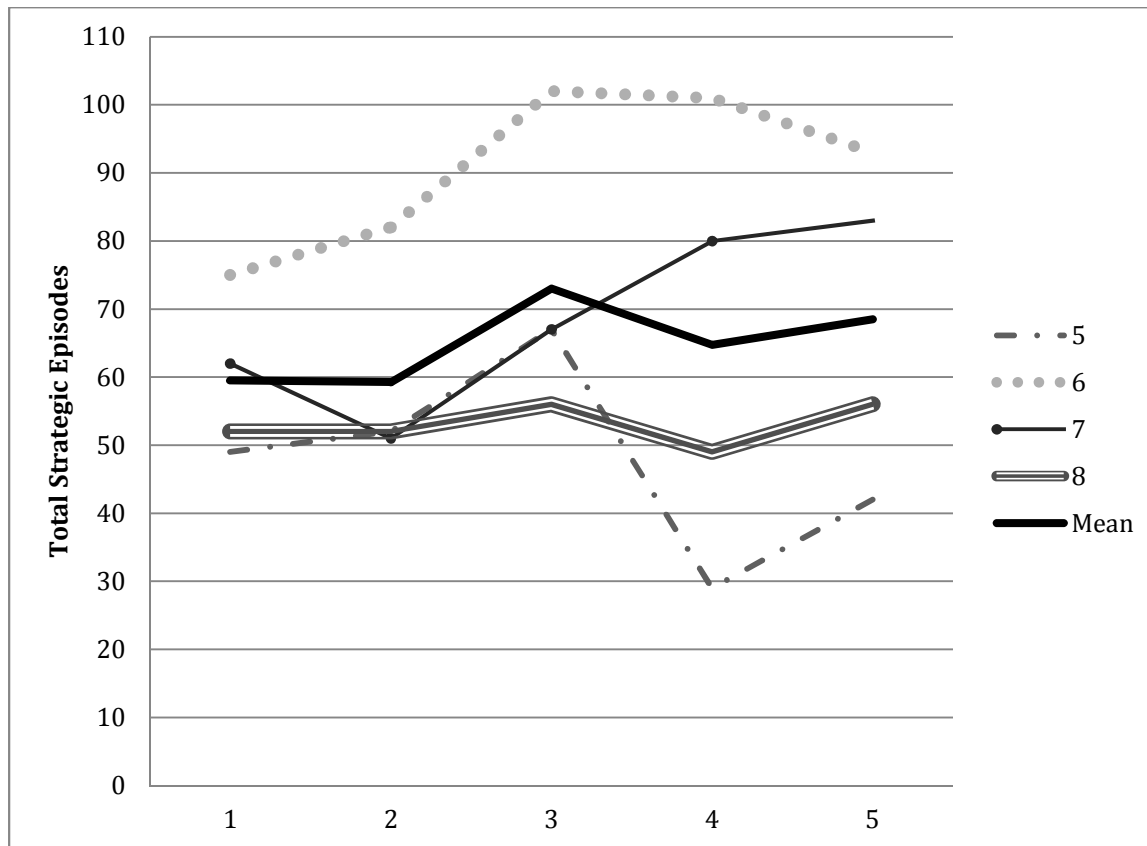


Figure 2. Total strategic processing episodes over five sessions for treatment dyads (n = 4).

**Comparison of ranges of difference scores of total strategy use.** To test this interpretation, I compared ranges of difference scores for both groups at all points in the study using the Moses Test of Extreme Reaction. Difference scores (see Table 8) were calculated by subtracting the total frequency count of all strategies applied at one session (e.g., Practice Session 1) from the total frequency count of all strategies applied at the previous session (e.g., Pretest Total). Although the ranges for difference scores between treatment and control groups did not differ statistically for *Practice Session 1-Pretest* or for *Practice Session 3-Practice Session 2*, results of this test indicated that the range of Change values in the control and treatment conditions differed at two points: *Practice 2-Practice 1* and *Posttest -Practice 3* ( $p < .001$  for both), with a

higher probability of extreme range values in the control condition. With mixed results, these data offer no conclusive evidence that would enable claims about the constraints that LINKS may have had on total number of strategic processes used. Consistent with the descriptive trend for the treatment dyads to use more strategies at practice sessions 2 and 3, however, there is some evidence to suggest that the intervention may have increased the total number of strategies that dyads used, while at the same time leveling the range of frequencies among treatment dyads in ways that the control condition did not experience.

Table 8

*Difference Scores for All Dyads*

Dyad (C,T)	P1-PRETEST	P2-P1	P3-P2	POSTTEST-P3
1 C	+5	+2	-31	+14
2 C	-21	+2	+0	+37
3 C	-32	+21	-24	+25
4 C	+8	-17	+26	-13
5 T	+3	+15	-38	+13
6 T	+7	+20	-1	-8
7 T	-11	+16	+13	+3
8 T	0	+4	-7	+7

*Note:* C= Control Dyad; T = Treatment Dyad; P1=Practice Session 1, P2 = Practice Session 2; P3 = Practice Session 3.

**Frequencies comparisons by strategy.** Although the first macro-level test of total frequencies on all strategies revealed no statistically significant differences between pretest and posttest for the two conditions, I also examined point-by-point differences for all strategies as a way to determine whether frequency of specific strategic episodes was influenced by the LINKS treatment. This could be considered a micro level of analysis of frequencies.



Tables 9 and 10 show mean frequency counts for treatment and control groups on all ten  $[(PST)^2 + (iC^3)]$  strategies at pretest and posttest.

Table 9

*Pretest and Posttest Mean Strategic Episodes  $[(PST)^2 + (iC^3)]$  for Control Group*

	A	B	C	D	E	F	G	H	I	J
Pretest	1.25 (.96)	.50 (.58)	<b>8.25</b> <b>(3.78)</b>	<b>14.5</b> <b>(2.88)</b>	.50 (.58)	.75 (.96)	<b>18.25</b> <b>(6.80)</b>	8.00 (6.48)	3.25 (1.70)	6.00 (3.74)
Posttest	1.75 (1.5)	.13 (.25)	<b>11.81</b> <b>(3.8)</b>	<b>15.31</b> <b>(.47)</b>	.25 (.50)	1.06 (1.36)	<b>17.43</b> <b>(6.24)</b>	5.00 (3.65)	1.93 (1.64)	6.94 (3.47)
Change	.50 (.54)	.37 (-.33)	<u>3.56</u> (.02)	.81 (-2.41)	-.25 (-.08)	.31 (.40)	-.82 (-.56)	<u>-3.00</u> (-2.83)	<u>(-1.32)</u> (-.06)	.94 (-.27)

*Note.* Mean and (Standard Deviation) reported in each cell. Change = Change in Mean (change in Standard Deviation) from pretest to posttest. A: Purpose, B: Pre-existing Knowledge, C: Search, D: Source Selection, E: Type, F: Trustworthiness, G: Identify Important Information, H: Compare to pre-existing knowledge, I: Connect to other texts, J: Continually Update; **Bold** is used to highlight the three most frequently applied strategies. *Italics* are used to indicate greatest three mean changes from pretest to posttest.

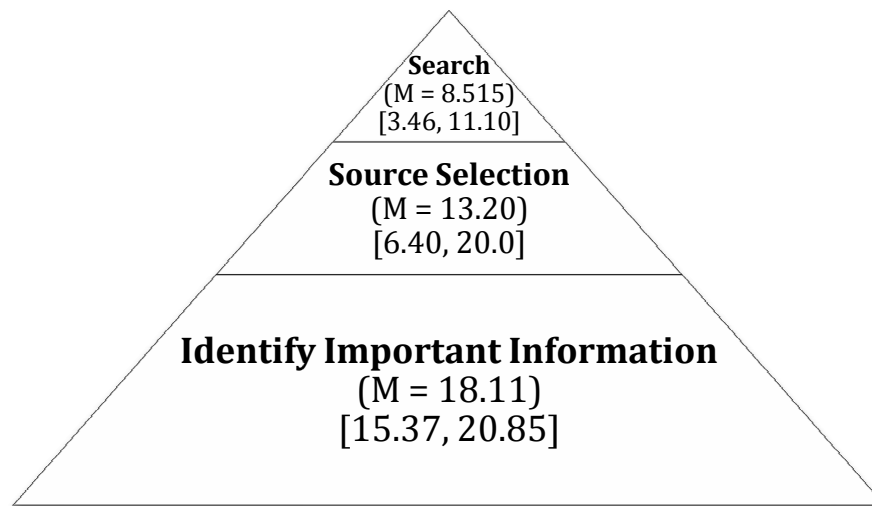
Table 10

*Pretest and Posttest Mean Strategic Episodes  $[(PST)^2 + (iC^3)]$  for Treatment Group*

	A	B	C	D	E	F	G	H	I	J
Pretest	.50 (.58)	0.00 (0.00)	5.50 (3.51)	<b>8.00</b> <b>(4.24)</b>	0.00 (0.00)	.50 (.58)	<b>20.00</b> <b>(4.16)</b>	5.25 (4.72)	1.75 (1.5)	<b>8.25*</b> <b>(3.78)</b>
Posttest	.50 (.58)	1.75 (.96)	<b>8.50</b> <b>(5.97)</b>	<b>15.00</b> <b>(9.41)</b>	2.50 (2.52)	3.25 (2.50)	<b>16.75</b> <b>(4.5)</b>	4.25 (.96)	3.00 (2.16)	7.25 (3.30)
Change	0.00 (0.00)	1.75 (.96)	<u>3.00</u> (2.46)	<u>7.00</u> (5.17)	2.50 (2.52)	2.75 (1.92)	<u>-3.25</u> (.34)	-1.00 (-3.76)	1.25 (.66)	-1.00 (-.58)

*Note.* See Table 8 for description of cell contents and code abbreviations. \*At pretest, the mean for J: *Continually Updating* was slightly higher than for D: Source Selection but at posttest the pattern of search, source selection and identify important information was consistent between treatment and control groups with updating being the fourth most frequently applied strategy. **Bold** is used to highlight the three most frequently applied strategies. *Italics* are used to indicate greatest three mean changes from pretest to posttest. All values rounded to nearest hundredth when value in thousandths place is equal to or more than .005.

***Descriptive analyses of strategy by strategy frequencies data.*** First, I explored the data visually. From these two tables, it is clear that *Identify Important Information* (G) was the most common strategic activity for both groups. This means that for all dyads, the most frequent strategic activity at pretest and posttest was the identification of topic-relevant information in individual texts. *Source Selection* (D) was the second most frequently applied strategy at pretest and posttest. *Search* (C) was the third most frequently applied strategy for the control group at pretest and posttest. For the treatment group, Search was the fourth most frequently applied strategy at pretest (only marginally behind continually update) and third most frequently applied at posttest. Generally, these data suggest that for all dyads, at pretest and posttest, irrespective of treatment received, the general frequency structure of these three processes looks like a pyramid (see Figure 3) with Search at the top, Source Selection in the middle, and Identify Important Information at the bottom as the foundation.



*Figure 3.* Pyramid representation of search, source selection, and identify important information (n=8) with mean frequency (M) and [95% confidence intervals] reported.

These frequencies are consistent with what we might expect based on the nature of the inquiry task itself, and also the logical sequence of tasks that lead to information gathering and then integration. Building an integrated understanding of a topic requires that there be information to integrate, so the task itself may have driven participants to focus most frequently on the identification of important, or topic-relevant information. This finding, in particular, is consistent with other studies that have found children and adolescents prioritize content relevance over other factors when considering the value of multiple texts, particularly as they begin their inquiry processes (Braasch, et al, 2009; Hirsch, 1999). Further, it makes sense that source selection was more frequently applied than search because students did return to the same search engine results page (SERP) repeatedly to find another text for close reading. Search, of course, is a precursor to source selection and to the identification of important information. These data show that these participants tried to identify important information in the texts they chose to read more frequently than searching on new keyword phrases or selecting new texts and that this trend was not influenced by the LINKS treatment relative to the control experience.

Importantly, the fourth and fifth most frequently applied strategic episodes for both groups at posttest were *Continually Update (J)* and *Compare to Background Knowledge (H)*. Again, the consistencies here suggest no real impact of intervention over the effect of the larger task at hand, but suggest a consistent turn to “self” as a strategy for constructing understanding across multiple texts for both groups.

If we consider the magnitude and directionality of pretest to posttest change for both groups (summarized in Tables 9 and 10) we see that for six of the 10 strategies, the treatment condition demonstrated a larger change from the pretest values. For strategy *H: Compare to Background Knowledge*, both groups decreased their use, but the control condition showed a

greater decrease [ $\downarrow\downarrow$ ]. For Strategy *Connect to other texts (I)*, the treatment group increased its use but the control condition decreased its use by a larger margin [ $\uparrow\downarrow$ ]. For strategy *Continually Update (J)*, the control condition increased their application slightly, but the posttest level for the treatment condition remained higher overall, even though the score decreased by 1.00 [ $\uparrow\downarrow$ ]. Interestingly, the largest change for the treatment condition between pretest and posttest (+7.0) was on *Source Selection (D)*. Control participants increased very little on this strategy (+.81). Treatment participants also decreased their use of Identify Important Information from pretest to posttest by a larger margin (-3.25) than the control condition (-.82). These data may suggest that LINKS supported a turn toward active selection of texts (i.e., source selection) vs. closer reading of individual texts (i.e., identify important information). Together, these descriptive frequencies by strategy, suggest greater change for the treatment condition from pretest levels, even though the general structure of the first, second and third most frequently applied strategies were consistent between groups at posttest.

***Strategy by strategy frequencies comparisons.*** To determine whether treatment and control groups differed on any frequencies of strategy use, at any point in the study, a series of non-parametric Mann-Whitney U tests were conducted because assumptions for normality and homogeneity of variance were not consistently met for all data. This test is appropriate for comparison of differences between two conditions when different participants have been used in the conditions (Field, 2009, p. 540). Using a series of tests like this increases the probability of Type I error by capitalizing on chance, but I used this approach to examine trends or indication of disruption in frequencies applications that could be triangulated with other analyses. Statistically significant results are interpreted cautiously.

All null hypotheses for between groups differences were retained for frequencies

comparisons of *Purpose (A)*, *Search (C)*, *Source Selection (D)*, *Identify Important Information (G)*, *Compare (H)*, *Connect (I)*, *Continually Update (J)*, *Construct Meaning in a Single Text (M)* and *Trustworthiness During Reading (Y)* episodes at all five points in the study. For these variables, we conclude that the population distributions from which the control and treatment data were drawn are the same. There was, however, sufficient evidence to reject the null hypothesis for three strategic episode codes at certain moments in the study.

At posttest, the frequency distributions of control and treatment groups on the Pre-Existing Knowledge (B) code were found to differ statistically significantly  $U=0.00$ ,  $Z = 2.381$ ,  $p = .029$ ,  $r = .84$ . The U-value of zero indicates that all ranks in the control group were lower than the ranks of the treatment condition. The effect size,  $r$ , (calculated using the formula  $r = Z/\sqrt{N}$ ) is large (Cohen, 1992, p. 157). This suggests that treatment participants continued to discuss their pre-existing knowledge at posttest as they had been instructed to do during practice.

Group distributions also differed statistically on frequency of *Type (E)* codes at practice session 2,  $U = 0$ ,  $Z = 2.53$ ,  $p = .029$ ,  $r = .89$ , and *Trustworthiness (F)* codes at practice session 2,  $U = 0$ ,  $Z = 2.477$ ,  $p = .029$ ,  $r = .88$  and practice session 3  $U = 0$ ,  $Z = 2.366$ ,  $p = .029$ ,  $r = .84$ . This is logical because participants were prompted to evaluate texts critically as part of the LINKS intervention.

For each statistically significant difference observed through these analyses, the mean frequency for the treatment group was higher than the mean for the control group. Although the total number of strategic episodes may not have differed between pretest and posttest for treatment and control groups, these analyses suggest that LINKS may have enabled increased evaluative processing as the intervention was administered, though not at posttest. At posttest, treatment participants did seem to discuss their background knowledge more frequently than

their peers who received the control experience.

***Comparisons of strategy by strategy relative frequencies.*** It was also assumed that differences within or between groups in *relative frequency* of strategy use would indicate the choice, or strategic application of certain processes relative to other strategic processing options. Goldman et al. (2012) used analyses of relative frequencies to explore “patterns” of strategy use (p. 366) because relative frequency is a proportion, and comparisons therefore account for the raw total differences in strategy use among dyads.

Table 11 shows *relative frequencies* for control participants on all  $[(PST)^2 + (iC)^3]$  strategy codes at pretest and posttest. Table 12 shows relative frequencies for treatment participants on the same strategy codes at pretest and posttest. I calculated relative frequency of strategy application by dividing per-code processing events by the sum total of processing events for each dyad, and for each of the five sessions in the study (though only pretest and posttest data are presented in these tables).

Consistent with my analyses for frequencies data, I also used the Mann-Whitney U test to compare distributions between groups for all relative frequencies at each point in the study. Results of these analyses were nearly identical. Null hypotheses were retained for all comparisons of *Purpose (A)*, *Search (C)*, *Source Selection (D)*, *Identify Important Information (G)*, *Compare (H)*, *Connect (I)*, *Continually Update* and for the added codes of *Construct Meaning within a Text (M)* and *Trustworthiness During Reading (Y)*, suggesting that the distributions for relative application of these strategies did not differ between treatment and control dyads at any of the five sessions in the study.

Null hypotheses were rejected, however for three categories of strategic episodes: (a) *Pre-existing Knowledge (B)* episodes at Posttest  $U = 0.00$ ,  $p = .029$ ,  $Z = 2.36$ ,  $r = .83$ , (b) *Type*

(E) episodes at practice session 1,  $U = 0.00$ ,  $p = .029$ ,  $Z = 2.38$ ,  $r = .84$ , practice session 2,  $U = 0.00$ ,  $p = .029$ ,  $Z = 2.46$ ,  $p = .029$ ,  $r = .87$  and practice session 3,  $U = 0.00$ ,  $p = .029$ ,  $Z = 2.32$ ,  $r = .82$  and (c) for *Trustworthiness (F)* episodes at practice sessions 2  $U = 0.00$ ,  $p = 0.029$ ,  $Z = 2.477$ ,  $r = .87$  and 3  $U = 0.00$ ,  $p = 0.029$ ,  $Z = 2.366$ ,  $r = .83$ . For each statistically significant comparison, the mean relative frequency for the treatment condition was higher than for the control condition. These results suggest that the LINKS intervention may have changed particular patterns of *Type (E)* and *Trustworthiness (F)* activities in the treatment condition during practice, but that at posttest, the effect did not transfer. Importantly, the relative application of *Pre-existing Knowledge (B)* strategies in the treatment condition was statistically significant at posttest, suggesting that without instructional prompting, students did transfer the practice of thinking about background knowledge before Search to the posttest context, and that this was not a strategic process applied by the control condition to the same extent.

Table 11

*Comparison of Pretest and Posttest Mean Relative Frequencies  $[(PST)^2 + (iC^3)]$  for Control Group*

	A	B	C	D	E	F	G	H	I	J
Pretest	.02 (.01)	.01 (.01)	<b>.14</b> (.10)	<b>.23</b> (.10)	.01 (.01)	.01 (.01)	<b>.26</b> (.04)	.10 (.08)	.05 (.02)	.08 (.04)
Posttest	.02 (.02)	.002 (.005)	<b>.17</b> (.03)	<b>.24</b> (.07)	.004 (.01)	.01 (.01)	<b>.26</b> (.07)	.07 (.04)	.03 (.02)	.10 (.03)
Change	(0) (+.01)	-.008 (-.005)	<b>+.03</b> ( <b>-.07</b> )	+.01 (-.03)	-.006 (0)	(0) (0)	(0) (+.03)	<b>-.03</b> ( <b>-.04</b> )	-.02 (0)	<b>+.02</b> ( <b>-.01</b> )

*Note.* Mean and (Standard Deviation) reported in each cell. Change = Change in Mean (change in Standard Deviation) from pretest to posttest. A: Purpose, B: Pre-existing Knowledge, C: Search, D: Source Selection, E: Type, F: Trustworthiness, G: Identify Important Information, H: Compare to pre-existing knowledge, I: Connect to other texts, J: Continually Update. **Bold:** Indicates largest proportions (relative frequencies) and largest changes. All values rounded to nearest hundredth when value in thousandths place is equal to or more than .005.

Table 12

*Comparison Of Pretest and Posttest Mean Relative Frequencies  $[(PST)^2 + (iC^3)]$  for Treatment Group*

	A	B	C	D	E	F	G	H	I	J
Pretest	.01 (.01)	.00 (.00)	<b>.08</b> (.05)	<b>.13</b> (.05)	0.00 (0.00)	.01 (.01)	<b>.35</b> (.11)	.08 (.07)	.03 (.03)	.13 (.04)
Posttest	.01 (.01)	.03 (.01)	<b>.12</b> (.05)	<b>.20</b> (.07)	.03 (.03)	.05 (.03)	<b>.26</b> (.07)	.07 (.02)	.05 (.03)	.11 (.03)
Change	0 (0)	+.03 (+.01)	<b>+.04</b> (0)	<b>+.07</b> (+.02)	+.03 (+.03)	<b>+.04</b> (+.02)	<b>-.09</b> (-.02)	-.01 (-.05)	+.02 (0)	-.02 (-.01)

*Note.* See Table 9 for description of table contents and meaning of codes. All values rounded to nearest hundredth when value in thousandths place is equal to or more than .005. **Bold** indicates largest proportions (relative frequencies) and largest changes.

***Within-groups analysis of change in Type (E).*** I followed these point-by-point comparisons with analyses of within-group change over time for both conditions in order to understand whether, or to what degree the conditions, themselves, led to change in relative frequency of *Type (E)* episodes for each group within the time frame of the study. This takes the analyses to a very specific micro level by focusing on the trajectory of the strategy that did differ between groups. Table 13 provides all mean and standard deviation data. Figure 4 shows the difference in relative application of *Type (E)* strategies over time for both the treatment and control conditions. It is clear that the treatment condition used more *Type (E)* strategies as a group than the control condition but the data did not meet all assumptions for repeated measures ANOVA. The assumption of sphericity was met  $W=.022, p=.067$  but the assumptions of homogeneity of error variances and normality were not. Given these violations, I used two non-parametric Friedman's tests to determine change over time for both groups. This test can be used to test differences across multiple sessions when the same participants have been used in each session (Field, 2009, p. 573). For the control group,  $\chi^2(4) = 3.895, p = .420$ , suggesting no statistically significant



change over the five sessions of the study. For the treatment condition,  $\chi^2(4) = 8.456, p = 0.076$  which is significant at the  $\alpha = .10$  level, but not at the .05 level. Despite the statistically significant between-groups differences at practice sessions 2, 3 and 4, the null hypotheses for within groups change over all five sessions were retained for both groups. The change within groups was not as substantial as the differences measured between them at certain points in the study. Importantly, for the treatment condition the most substantive change occurred at practice session 1 – the moment that the LINKS intervention was introduced. Although I conclude from a Wilcoxon Signed-Rank Test that there is insufficient evidence to conclude that relative frequencies of type episodes at pre-test and Practice Session 1 differ statistically [ $Z = 1.841, p = .066$ ] it is also clear from these data that episodes of *Type (E)* strategy application increased for the treatment condition when the intervention was introduced and remained higher than pretest for the remainder of the study.

Table 13

*Mean Relative Frequencies of Type (E) Episodes by Condition*

Session	Control M (SD)	Treatment M (SD)
1 Pretest	.01(.008)	.00 (.000)
2 Practice 1	.01 (.009)*	.03 (.019)*
3 Practice 2	.00 (.000)*	.03 (.023)*
4 Practice 3	.01 (.011)*	.04 (.006)*
5 Posttest	.004 (.008)	.04 (.031)

*Note:* All values rounded to nearest hundredth when the value in the thousandths place is .005 or higher. \* Statistically significant between-groups differences at these moments in the study.

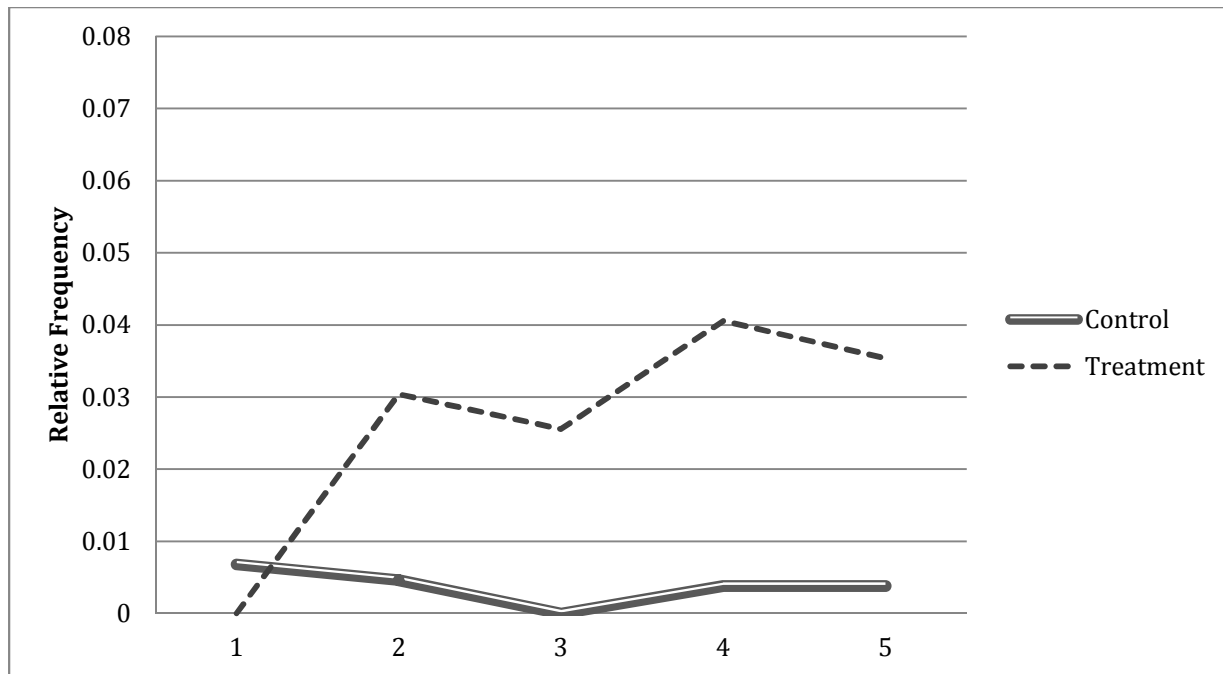


Figure 4. Mean relative frequency of Type (E) episodes over five sessions of the study for treatment and control groups.

**Within-groups analyses of change in Trustworthiness (*F*).** From Figure 5, it is evident that the relative frequency of *Trustworthiness (F)* episodes increased for the treatment condition, but stayed relatively flat for the control condition during the study, suggesting an effect of treatment on application of this strategy. Comparison of mean proportions of trustworthiness episodes using Repeated Measures ANOVA was not justified because the distributions for several mean values contradicted the assumption of normality.

I conducted non-parametric Friedmans' tests to compare repeated, related measures. For the control group, the result of the Friedman's test  $\chi^2(4) = 7.097, p = .131$ , indicated that control dyads did not change the relative frequency of their use of trustworthiness strategies over the five sessions of the study. For the treatment condition, the Friedman's test result  $\chi^2(4) = 9.975, p = 0.041$  was statistically significant with an effect size, reported as Kendall's  $w = .623$ . This suggests that the LINKS treatment had a statistically significant impact on the relative

application of trustworthiness strategies over the five sessions of the study and that the effect of the treatment was quite strong (Cohen, 1992).

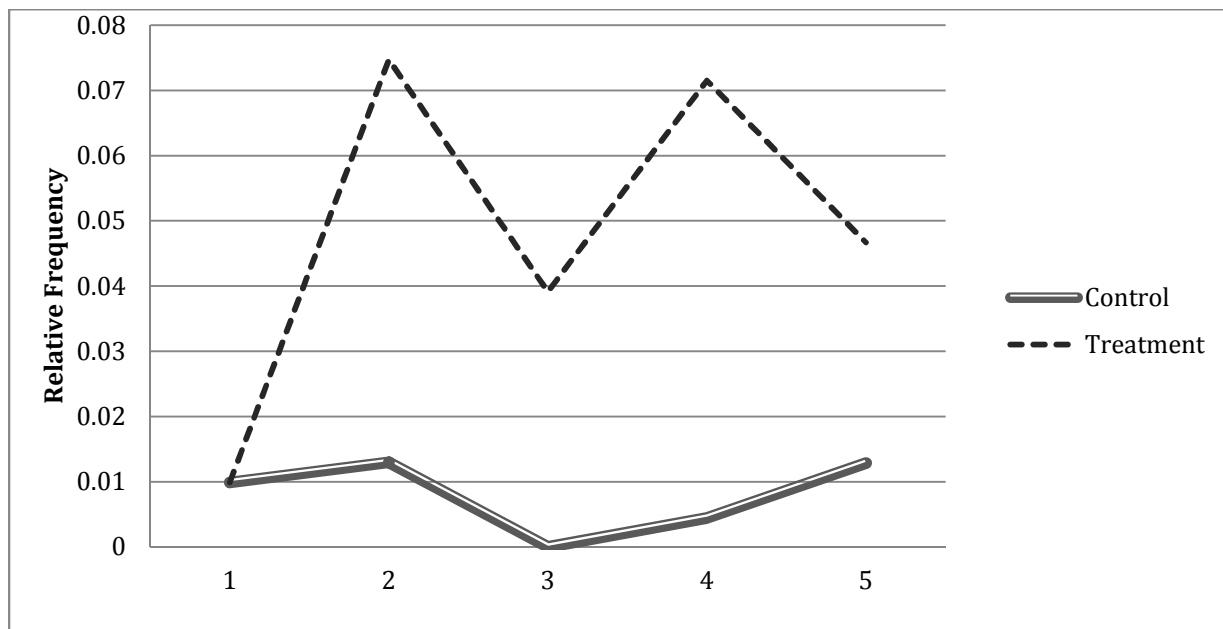


Figure 5. Mean relative frequency of Trustworthiness (F) episodes over five sessions of the study for treatment and control groups.

Table 14

*Mean Relative Frequencies of Trustworthiness (F) Episodes by Condition*

Session	Control M (SD)	Treatment M (SD)
1 Pretest	.01 (.01)	.01(.01)
2 Practice 1	.01 (.01)	.07 (.04)
3 Practice 2	.00 (.00)*	.04 (.01)*
4 Practice 3	.004 (.01)*	.07 (.03)*
5 Posttest	.01 (.01)	.05 (.03)

*Note:* All values rounded to nearest hundredth when value in thousandths place is more than or equal to .005. \*Statistically significant mean differences.

**Analyses of time patterns.** A third way to determine the impact of the treatment condition on strategic episode application was to examine differences in students' use of time at pretest and posttest. I compared two metrics here. The first, *relative duration*, indicates the proportion of time that dyads spent using each strategy. Like relative frequency, it accounts for the total amount of time per session when, as in this study, the end time varies slightly (Bakeman & Quera, 2011, p. 98). For each group, I also analyzed *mean gap*, which is an average of the elapsed time between uses of the same strategy over a given session. Changes in relative duration and mean gap suggest change in patterns of strategy application—using certain strategies for longer or shorter periods of time or in quicker or less quick succession. To calculate these data, I used onset sequence data collected in Morae (Techsmith, 2012) for each code (e.g., 2:24.2, C (Search); 2:27.4, D (Source Selection). The difference between the onsets of two codes in sequence is taken as the *duration* of the preceding code.

First, I present findings from a series of statistical analyses. Then, I describe the series of graphs that demonstrate the structure of these data. Pretest and Posttest graphs for both groups are presented in sequence for relative duration and mean gap. I did not include data for practice sessions because I was looking to understand impact as measured at posttest relative to pretest.

To compare pretest and posttest results for treatment and control groups on relative duration and mean gap for each strategic process of  $[(PST)^2 + (iC^3)]$  and also codes *M* *Constructing Understanding in a Single Text* (*M*), *Notetaking* (*N*) and *Trustworthiness During Reading* (*Y*), I used a series of Wilcoxon matched-pair signed rank tests. This test is appropriate for comparing two measures of the same variable at two moments in time when the participants are the same in both measures (Field, 2009). I included *Notetaking* (*N*) in these analyses because some dyads seemed to spend a lot of time taking notes and I wondered if the treatment

conditions might have had an impact on the relative duration of note taking between and/or within groups. *Notetaking* (*N*) was not a strategy included in the  $[(PST)^2 + (iC^3)]$  framework, but as noted in the methods sections, students took notes using different colored pens to record information as they read online.

For the treatment condition, no pre-post comparisons were statistically significant at the .05 level of alpha for relative duration and mean gap. This means that LINKS did not have a statistically significant impact on the relative duration of any coded strategic activity or on the mean gap for any strategy code. The same finding was true for the control condition. From pretest to posttest, no Wilcoxon matched-pair signed rank test was statistically significant at the .05 level of alpha.

To examine between-groups differences, I used a series of Mann-Whitney U tests. At pre-test, the only statistically significant difference between groups was on strategy *Source Selection* (*D*)  $U = .500, Z = -2.191, p = .029, r = .77$  with the control group having a higher mean value. At posttest, the difference between groups was not statistically significant  $U = 5.50, Z = -.726, p = .49$ .

***Descriptive analysis of time patterns.*** Descriptively the graphs and mean values for treatment and control groups do suggest slightly different trends in the patterns of relative duration and mean gap that may warrant further investigation. Figures 6, 7, 8, and 9 show pretest to posttest change for both groups on *relative duration*. Figures 10, 11, 12, and 13 show pretest to posttest change for both groups on *mean gap*. Data are also summarized in Table 15.

***Comparisons of relative duration.*** As is visible in Figures 8 and 9, for the treatment condition, relative duration for strategies *Identify Important Information* (*G*), *Continually Update* (*J*), *Construct Meaning in Single Text* (*M*) and *Notetaking* (*N*) decreased from pretest to posttest,

suggesting that treatment participants spent relatively less time identifying information in texts, explicitly updating their understanding, constructing understanding in single texts and taking notes (see Table 14). Relative duration increased, however, for the treatment condition on *Search (C)*, *Source Selection (D)*, *Type (E)* and *Trustworthy (F)* codes. Strategies *Connect to Other Texts (I)* and *Compare to Background Knowledge (H)* changed by only one percent up and down respectively.

Like the treatment group, the control condition (see Figures 6 and 7) also engaged relatively less time on *Notetaking (N)* at posttest and increased their engagement of *Type (E)* and *Trustworthy (F)* codes. Although the mean scores at posttest look nearly the same for both groups, it is interesting to note that for *Search (C)* and *Source Selection (D)*, the control condition spent relatively *less* time on these two strategic activities at posttest whereas the treatment condition increased the relative amount of time spent on them. On the other hand, the control condition spent relatively *more* time on strategies *Identify Important Information (G)*, *Continually Update (J)* and *Construct Meaning in Single Texts (M)*, as the treatment condition engaged these strategies *less*. Although differences were not statistically significant, the trend for treatment and control groups to move in opposite directions on these two sets of strategies (C & D ; G, J & M) may be worthy of further investigation. One question raised by this descriptive trend would be whether the LINKS intervention expands the range of strategies to which students have access, so that they spend less time on any given text (i.e., strategies *G (Identify Important Information)* and *M (Construct Meaning from a Single Text)*) and more time engaged in search and source selection which could, in the end, enable the integration of more texts and potentially more trustworthy texts too.

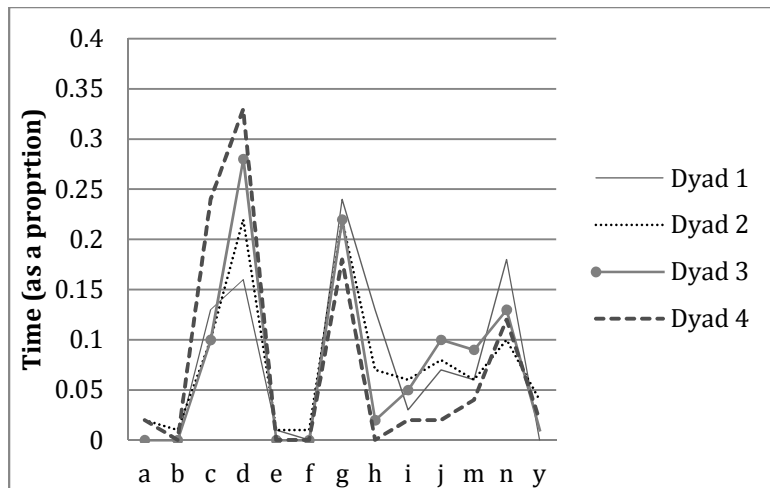


Figure 6. Control group: Pretest relative duration.

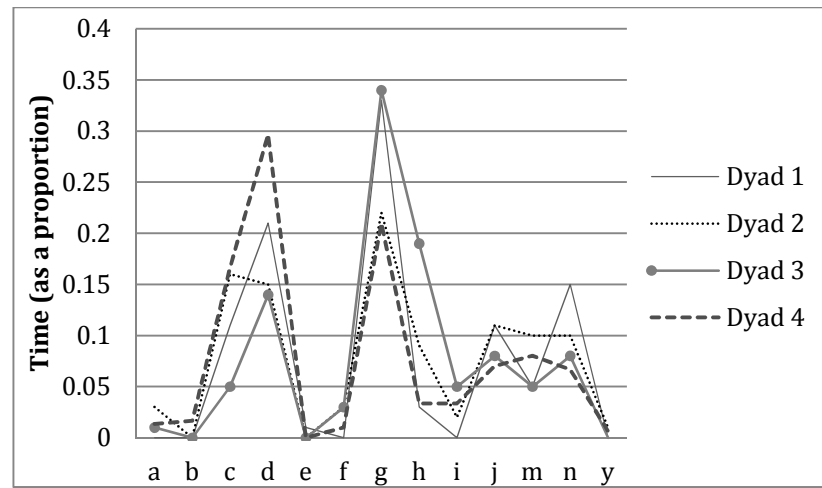


Figure 7. Control group: Posttest relative duration.

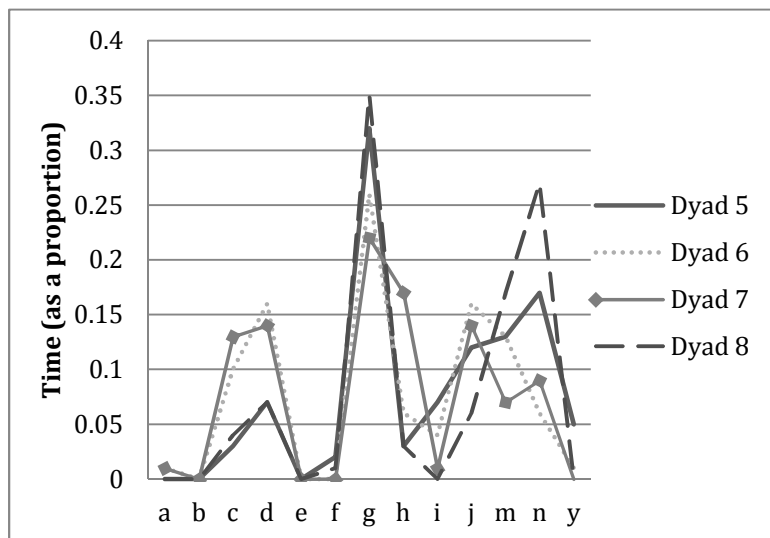


Figure 8. Treatment group: Pretest relative duration.

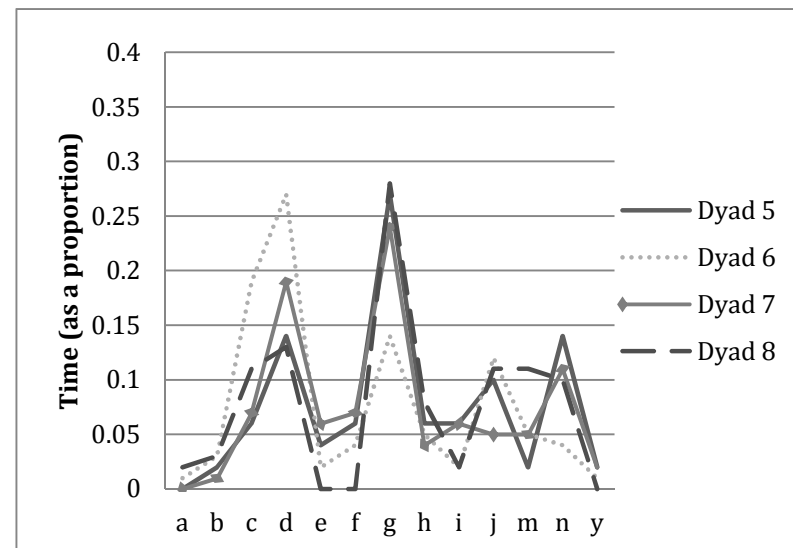


Figure 9. Treatment group: Posttest relative duration.

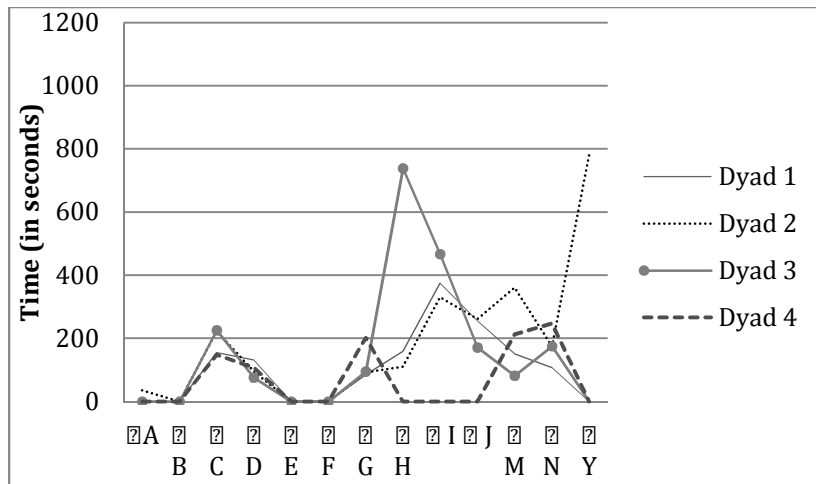


Figure 10. Control group: Pretest mean gap.

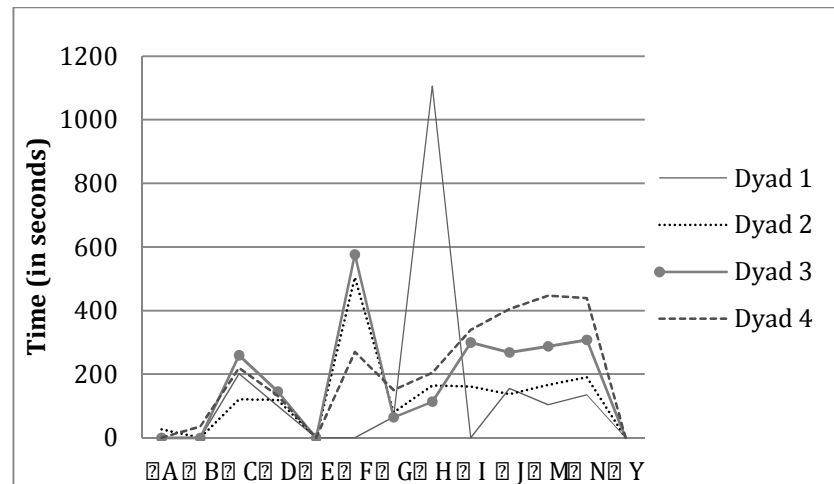


Figure 11. Control group: Posttest mean gap.

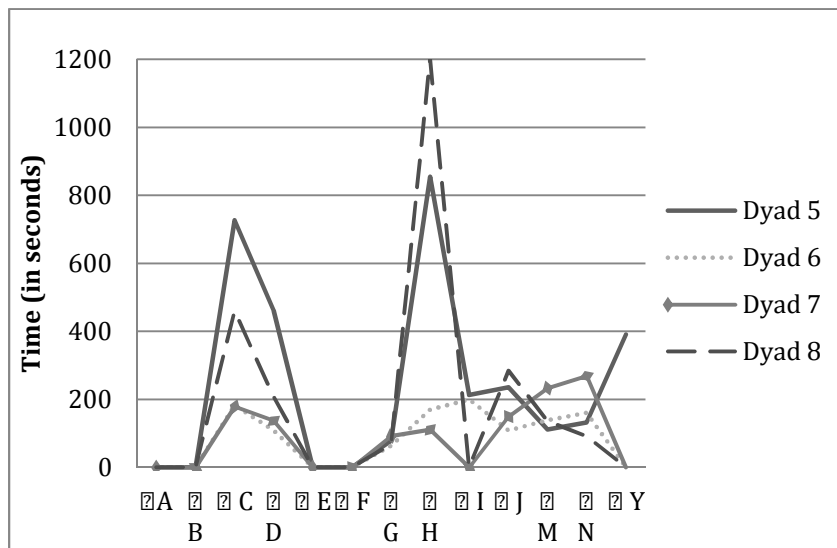


Figure 12. Treatment group: Pretest mean gap.

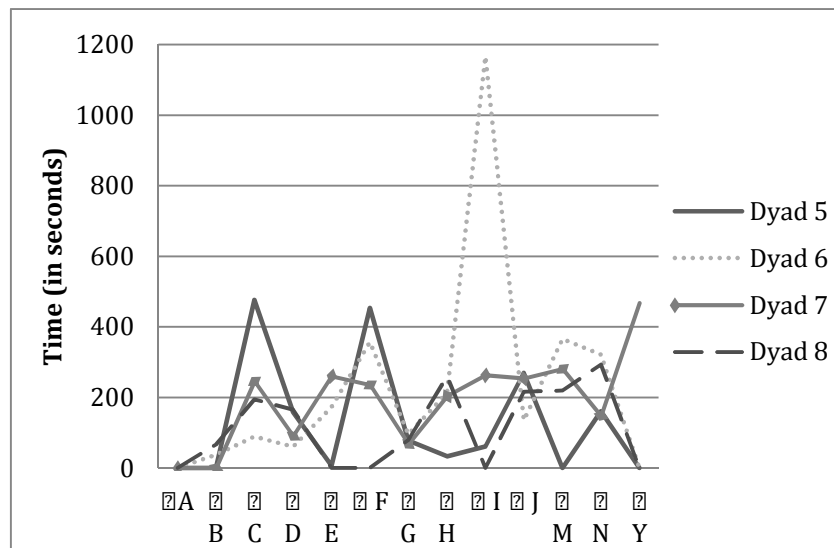


Figure 13. Treatment group: Posttest mean gap.



**Comparisons of mean gap.** The mean gap for most treatment dyads appeared to decrease for the strategy *Compare to Background Knowledge (H)*. The average mean gap value at pretest for the treatment group on *Compare to Background Knowledge (H)* was 584.25 seconds (SD = 532.00). It decreased to 177.19 seconds (SD = 99.58) at posttest, suggesting that these dyads did engage background knowledge in quicker succession during online inquiry at posttest. Likewise, for strategies *E (Type)* and *F (Trustworthy)* the pre-test values were effectively zero for the treatment group. At posttest, mean gap values were quite variable, but measureable. The posttest average for *Type (E)* = 109.84 (SD = 128.31) and for *Trustworthiness (F)* = 337.96 (SD = 260.62). Although within-groups and between groups differences were not statistically significant on any code, descriptive evidence suggests that LINKS may have disrupted mean gap for treatment *Type (E)*, *Trustworthy (F)* and *Compare to Background Knowledge (H)*-- were found to differ between groups on relative frequency during practice sessions.

Table 15

*Mean Values of Relative Duration for Control and Treatment Groups*

Strategy (Code)	Control		Treatment	
	Pretest	Posttest	Pretest	Posttest
Search (C)	.14 (.07)	.12 (.05) -	.07 (.05)	.11 (.06) +
Source Selection (D)	.25 (.07)	.20 (.07) -	.11 (.05)	.18 (.06) +
Type (E)	.005 (.006)	.025 (.005) +	0.00 (0)	.03 (.02) +
Trustworthy (F)	.003 (.005)	.017 (.015) +	.008 (.01)	.04 (.03) +
Identify Info (G)	.21 (.02)	.28 (.07) +	.29 (.06)	.23 (.06) -
Compare (H)	.05 (.05)	.09 (.02) +	.07 (.07)	.06 (.02) -
Connect (I)	.04 (.02)	.03 (.02) +	.03 (.03)	.04 (.02) +
Continually Update (J)	.07 (.03)	.09 (.02) +	.12 (.04)	.10 (.03) -
Single Text Meaning (M)	.06 (.02)	.07 (.02) +	.13 (.04)	.06 (.04) -
Notetaking (N)	.13 (.03)	.09 (.04) -	.15 (.09)	.10 (.04) -

Table 15 (cont'd)

*Note:* + indicates increase in relative duration from pretest to posttest, - indicates a decrease.

**Summary of section 1 results.** In response to the first research question about the impact, if any, of the LINKS treatment versus the control group experience on students' application of  $[(PST)^2 + (iC^3)]$  strategies, these results paint a picture of some complexity. In terms of total number of strategies applied, neither condition hastened change. Certainly there was variability among dyads at each session of the study, but overall, LINKS had no statistically significant impact on total frequency of strategy application at any point in the study between or within groups.

Analyses of the most frequently used strategies showed a common underlying structure that was not influenced by treatment. Data showed that the most frequently engaged activities are precisely those we might expect students to engage in order to complete the tasks, and in a predictable order. The most frequently applied strategy for both groups was *Identify Important Information (G)*. The second most frequently applied strategy was *Source Selection (D)*. The third most frequently applied strategy was *Search (C)*. Interestingly, the treatment group increased their use of *Source Selection (D)* and *Search (C)* by a larger margin than the control at posttest and decreased their application of *Identify Important Information (G)*, suggesting a treatment group shift in strategic focus toward more texts vs. a focus on gathering information in single texts.

Relative frequency of certain strategic episodes at certain moments in the study did differ between treatment and control conditions. Importantly, treatment participants engaged more *B (Pre-existing Knowledge)* episodes at posttest than the control group, suggesting that treatment students continued to use this strategy at posttest, even when the LINKS intervention was

removed. On the other hand, increased application of critical evaluations skills taught with LINKS did not persist at posttest. Results showed that the LINKS treatment enabled students to engage relatively more *Type (E)* and *Trustworthiness (F)* episodes than control participants during practice sessions but that this between-groups effect was lost at posttest when treatment participants engaged in their inquiry task without instructional support.

Interestingly, for *Trustworthiness (F)* episodes, a statistically significant within-group difference was found for the treatment condition. Over the span of the study, these four dyads did, therefore, engage this strategy relatively more frequently. Even though there was a within-group change for this strategy for the treatment condition, the control experience may have led to enough of a change in relative application of *Trustworthiness (F)* that the difference in mean values was insufficient to conclude the two groups were sampled from different populations at posttest.

For the participants in this study, this could mean a couple of things. First, it could mean that the LINKS intervention was effective in disrupting relative application of certain strategies while it was being administered (i.e., for *Type (E)* and *Trustworthiness (F)* episodes) but that the effect did not transfer to the posttest condition. It could also mean that these strategies, among all of the strategies included in  $[(PST)^2 + (iC^3)]$  and modeled for students during the LINKS intervention, are the most responsive to the LINKS instruction, but they may have been harder for this sample of adolescents to apply in the absence of guided supports.

Analyses of relative duration and mean gap data confirm that the LINKS treatment had no statistically significant impact on the percentage of time students apportioned to particular strategic activities, or to the rate with which they engaged them at posttest. Like the frequencies data that demonstrated a common pattern of strategy application in terms of frequencies between

treatment and control dyads (i.e., Identify Important Information, Source Selection, Search), these data showed a common structure in use of time. Even though descriptively, the treatment participants allotted more time at posttest to *Search*, *Source Selection* and to critical evaluation of texts than they had at pretest; and even though, statistically, the treatment groups explicitly engaged background knowledge more often than the control group at posttest, the general conclusion from section 1 is that LINKS enabled only targeted disruption in strategy use and did not enable any wholesale shifts in strategic processing between groups.

## **Section 2: LINKS' Impact on Trace Indicators of Integration in Students' Writing**

The second section examines the LINKS intervention's impact on trace indicators of integration in students' written arguments. Participants wrote independently for twenty minutes after reading with a partner for thirty minutes. Based on analyses of normality and homogeneity of variance, parametric methods were used to evaluate the overall quality of essays, as scored on the ACT rubric, but non-parametric tests were used to explore trace indicators of integration (TII) in the written products.

**Evaluation of pretest quality of written persuasive arguments.** Mean scores of essay quality at pretest, as measured on the ACT rubric, were compared for treatment and control groups using an independent samples t-test.

The Levene's test showed that variances between groups were homogenous  $F(1, 14) = 2.694, p = .123$ . The Shapiro-Wilk test showed that the distributions of each group could be assumed to be normally distributed  $w(\text{control}) = .875, p = .168$  and  $w(\text{treatment}) = .916, p = .397$ . The null hypothesis, which assumed identical mean scores on quality of persuasive arguments between groups, was retained,  $t = 1.169, p = .262$ . The overall quality of the persuasive arguments produced by control and treatment participants, as measured by mean scores on the

ACT rubric, therefore, were assumed to be from the same population.

Repeated measures ANOVA showed no statistically significant within-groups main effect of “session”  $F(2.780, 38.92) = 1.428, p = .250$  (with Huyn-Feldt correction because sphericity was not observed) or treatment condition  $F(1,14) = .577, p = .460$  on ACT rubric scores. For participants in this study, these findings suggest that quality of persuasive writing, as measured by the ACT rubric, was not influenced by repeated practice (control) or by the LINKS treatment.

**Comparison of trace indicators of integration in persuasive arguments.** Mean values for the trace indicators of integration index (TII Index) were also compared. Table 15 includes descriptive statistics for treatment and control groups for the five essays. Figure 14 displays the mean comparisons graphically.

Table 16

*Summary of Mean TII Scores for Control and Treatment Groups*

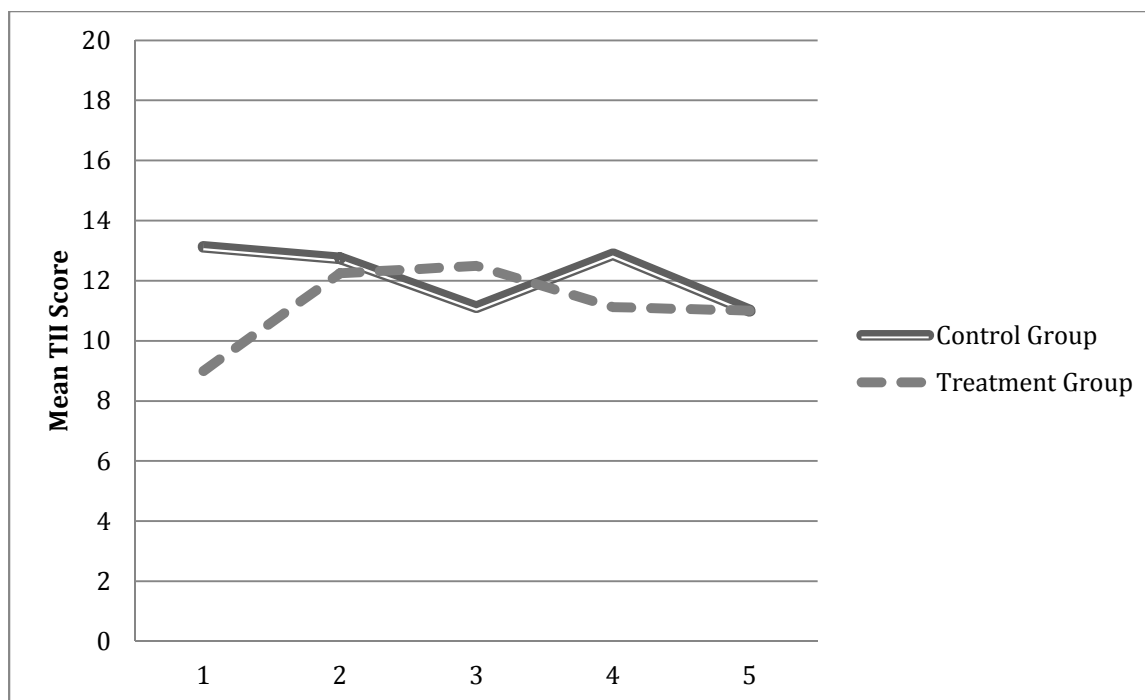
	Control		Treatment	
	M (SD)	95% CI	M (SD)	95% CI
1 Pretest	13.13 (1.73)	[11.93, 14.32]	9.00 (4.84)	[5.65, 12.35]
2 Practice 1	12.75 (2.66)	[10.93, 14.56]	12.25 (3.81)	[9.60, 12.90]
3 Practice 2	11.12 (3.31)	[8.82, 13.41]	12.50 (3.89)	[9.81, 15.20]
4 Practice 3	12.88 (3.09)**	[10.77, 14.99]	11.12 (2.99)	[9.05, 13.20]
5 Posttest	11.00 (2.39)**	[9.34, 12.66]	11.00 (3.42)	[8.62, 13.37]

\*\*Normality assumption violated.

The Shapiro-Wilks test confirmed the assumption of normality was met for all treatment distributions but not for control group essays at practice session 3 (essay 4) or at posttest (essay 5). Given these violations, I used non-parametric tests to compare between-group differences and

within group change over the course of the study on the TII measure. Descriptively, it is important to point out that the treatment condition mean score at pretest was lower than the score for the control condition. At posttest, the treatment condition's mean score increased and the control condition's mean score decreased. A between-groups comparison of pre-test mean scores using the Mann-Whitney U test was not statistically significant  $U = 15.00, Z = -1.85, p = .083$ . Although scores seemed to track in the expected direction for the treatment condition, no statistically significant between-groups results were found at any point in the study, including at posttest,  $U = 27.5, Z = -.483, p = .645$ .

Results of Friedman's ANOVA, which tests repeated-measures change within groups, were not statistically significant for control or treatment groups. For the control group,  $\chi^2(4) = 4.189, p = .381$ . For the treatment condition,  $\chi^2(4) = 7.709, p = .103$ . Given the increase in the mean TII scores seen at practice session 1, and then maintained by the treatment condition over the remainder of the study, I also examined the mean differences between pre-test and practice session 1 using the Wilcoxon signed-rank test. The result was not strictly significant at the alpha = .05 level,  $Z = -1.895, p = .058$ . However, the effect size,  $r = .67$  suggested an effect worthy of consideration. A Wilcoxon signed-rank test that compared the pre-test scores with scores at practice session 2 were, in fact, statistically significant,  $Z = -2.384, p = .017, r = .84$ .



*Figure 14.* Mean Trace Indicators of Integration Index scores for both groups on all five essays.

Given (a) the timing of this improvement, (b) the effect sizes, and (c) that the improvement was generally sustained over the remainder of the study for the treatment condition, I interpret this as an indicator of disruptive promise which nudged participants toward more integrative action. Compared to the control condition that did not see a similar increase at any moment in the study, LINKS seems to have shaken up these students' writing processes so that they were able to demonstrate more evidence of integrative thinking as measured on this rubric.

**Comparisons of discrete indicators of integration.** To further explore the impact of the intervention on trace indicators of integration, I conducted discrete pre-post non-parametric comparisons of specific items included in the integration rubric. These analyses allowed me to more closely examine the particular aspects of integration that may have been more or less influenced by the LINKS intervention compared to the control group, but also within each group.

For these pre-post analyses, I selected three items from the rubric that were most closely aligned with the items used to measure Synthesis in the ORCA-Open (Leu et al., 2012)—(recall that it uses evidence of intertextuality, and integration of details from two websites in a written product) but also, arguably, the most concrete or central indicators of integration of multiple texts across multiple information sources. Of the ten items included in the rubric, (a) the inclusion of information from more than one Internet text, (b) the use of corroborating information from two or more Internet texts, and (c) the use of counter-facts to the main argument that were collected from websites not used to inform the main argument were selected for discrete analysis. I also compared evidence of integration of background knowledge because treatment participants were instructed to talk about and write down their background knowledge as a part of LINKS. Given the importance of background knowledge as a schematic foundation for new understanding (e.g., Anderson & Pearson, 1984; Kintsch, 1998), and the statistically significant finding that the treatment condition did engage background knowledge more frequently than the control participants during their reading processes (see analyses of frequencies and relative frequencies above) I wondered whether treatment and control participants would differ in their use of background knowledge in their written arguments as well.

Results of a Mann-Whitney U test showed that at pretest, control and treatment groups seemed equally likely to include information from more than one Internet text in their written arguments  $U = 20.00, Z = -1.852, p = .064$ . The same was true at posttest  $U = 28.00, Z = -1.00, p = .317$ .

The groups were also equally likely to include corroborating information in their written arguments from two or more Internet texts at pretest,  $U = 26.00, Z = -.77, p = .441$ . They were



also equally likely to include corroborating facts from two or more texts at posttest  $U = 20.00$ ,  $Z = -1.852$ ,  $p = .064$ .

On their use of counter-facts to the main argument and the use of background knowledge in their written arguments, control and treatment groups were, however, found to differ at posttest. Specifically, the rubric accounted for the inclusion of counterpoints to the central argument collected from one or more sources that were *different* from the sources used to construct the central argument. In effect, this criterion was designed to tap into students' gathering of multiple perspectives from multiple texts and then whether that gathering resulted in the inclusion of multiple perspectives in the essay. On this criterion at pre-test, the control condition mean rank (10.56) was statistically significantly higher than the mean rank for the treatment condition (6.44)  $U = 15.5$ ,  $Z = -2.031$ ,  $p = .042$ , with an effect size  $r = .51$  meaning that the control participants were more likely to show evidence of this process in their essays at the start of the study, and that the size of that effect was large (Cohen, 1992). At posttest, however, the means were flipped. The mean rank for the treatment condition was 10.50 and for the control, it was 6.50 with  $U = 16$ ,  $Z = -1.936$ ,  $p = .053$ ,  $r = .48$ . Although this between-groups comparison was not strictly statistically significant at the .050 level of alpha, the size of the effect at posttest was large. Moreover, the within-group pre-post Wilcoxon Signed Ranks comparison for the treatment condition on this criterion was statistically significant  $Z = -2.236$ ,  $p = .025$ ,  $r = .79$ , suggesting that by posttest, treatment participants were able to include more counterpoints in their essays that they gathered from texts that were not also used to construct their main argument. Again, the calculated effect size was large for this pre-post difference (Cohen, 1992).

For the control condition, however, the pre-post Wilcoxon Signed Rank comparison revealed no statistically significant difference  $Z = -1.265$ ,  $p = .206$ . For the control group, it therefore cannot be said that the pretest and posttest scores were sampled from different populations. Together, these data suggest that more change occurred in the treatment condition on this criterion of “counterpoint use” than in the control condition.

It is important to note that for the control, there may have been a ceiling effect. They did start out with higher mean ranks than the treatment condition and may not have had as much room for improvement. That said, the treatment condition students did see significant gains on this factor, suggesting that for these students, LINKS may have enabled them to bring together more facts from more diverse perspectives in their written arguments.

Finally, treatment participants were also found to have integrated more evidence of background knowledge in their posttest written arguments than the control group,  $U = 11.5$ ,  $Z = -2.45$ ,  $p = .014$ ,  $r = .61$ . At pretest on this criterion, however, the groups were found to have been sampled from the same population,  $U = 28.00$ ,  $Z = -1.00$ ,  $p = .317$ . Moreover, the Wilcoxon Signed Rank comparison for pretest vs. posttest mean ranks revealed a statistically significant within-group difference for the treatment group,  $Z = -2.33$ ,  $p = .02$ ,  $r = .83$ , suggesting that at posttest, the treatment participants, who were found to make more explicit note of their prior knowledge on the topic while reading, also included that knowledge more often in their argumentative essays. The control condition did not change on this criterion between pretest and posttest,  $Z = 0.00$ ,  $p = 1.00$ . The implication and importance of these findings are summarized at the end of this section.

***Integration in students' written persuasive arguments.*** To provide a sense of what more and less integration looked like in students' written persuasive arguments, I provide two

contrasting examples: an essay that received a higher score (16/20) on the TII index vs. an essay that scored a lower score (8/20). Both essays were written on the same topic (Nuclear Power). The higher scoring essay was written by a control group participant at her first practice session. The lower scoring essay was written by a treatment group participant at her first practice session. The full evaluation rubrics for these essays are included in Appendix H, but annotations that document information sources or evidence of integration, collected from the participants' online inquiry sessions and notes, are provided in lists below each essay. I used color to indicate change of information source or evidence of a particular indicator of integration. (For interpretation of the references to color in this and all other essay annotations, the reader is referred to the electronic version of this dissertation.)

*Example 1: Higher Scoring Essay [16/20 on TII index] [Control participant, practice session 1]*

In light of ongoing conflict regarding nuclear power plants, many opinions are available and each a little bit different from the rest. For the most part, there are two main opinions regarding the use of nuclear power plants: they *should* be utilised and they *should not* be utilised. After conducting internet research, I would advise you to take full advantage of nuclear power. (1)

Although there are positive and negative effects of nuclear power- as there are pros and cons of any idea or action to be made-in this case, the positives outweigh the negative. (2) Among these positives are that we already know that nuclear power works- we have been using it since the 70's.(3) Nuclear power generation emits the least amount of carbon dioxide and greenhouse gases of all methods of energy generation. (4) It is easy to expand generation,(5) and in addition to this, expansion will create about 1,400-1,800 jobs to build each plant (with 400-

700 permanent jobs retained). (6) Also, nuclear power plants were designed to have long stretches of energy generation to minimize “down time” for refueling.(7)

A few negative ideas are joined at the hip to nuclear power: the threat of terrorism, and the haunting memories or accidents like Chernobyl and Three Mile Island.(8) Due to the radioactive nature of the materials involved in nuclear power plants, there is worry that terrorists will steal things like plutonium 239 or uranium and use them in weapons. (9) This is certainly a valid concern, but the plants would have security measures. (10) As for hesitation towards nuclear power because of previous accidents, all lives lost are tragedy, and it should not be taken lightly. That said, what do you have to do when you fall off your bike while learning to ride? You have to pick yourself up, dust yourself off, and try again. (11)

Nuclear power is the answer to a lot of our current energy needs. We need a clean source of reliable energy, something that we know will work for a long time and that won't run out. Nuclear power fits all of these criteria as well as holding massive potential for job growth.(12)

## Annotations

[based on inquiry session transcript, sites visited, texts read and notes recorded by the student]

1. Introduction that provides clear understanding of task prompt and expectations
2. Statement of position/thesis
3. not certain where she read this fact specifically, if at all, but her research about 3-mile island certainly informed an understanding of nuclear power use as early as the 1970s.
4. from [nuclearinfo.net](http://nuclearinfo.net) during the screencast

5. from the video -- the nuclear physicist who says this
6. from [nei.org/resourcesandstats/documentlibrary.../](http://nei.org/resourcesandstats/documentlibrary.../)
7. from the last site they read -- [www.fi.edu/guie/wester/benefits.html](http://www.fi.edu/guie/wester/benefits.html)
8. integrated ideas they read about at [world-nuclear.org/info/chernobyl/info7.html](http://world-nuclear.org/info/chernobyl/info7.html);  
[nuclearfissioni.wikispaces.com/cons+of+nuclear+fission](http://nuclearfissioni.wikispaces.com/cons+of+nuclear+fission)
9. [nuclearfissioni.wikispaces.com](http://nuclearfissioni.wikispaces.com) and from the video in the intro screencast
10. integrative inference -- her own insertion/response to the cons.
11. more integrative inferencing based on her own evaluation of the pros vs. cons
12. very integrative final statement that also provides a clear statement of position.

Notice how in this essay, the student included ideas extracted from several informational text sources, some of which were multimodal, used transitional phrases such as “that said” and “this is certainly a valid concern [sic], but...” that introduce the juxtaposition of contrasting viewpoints, and provided a clear statement of position informed by a synthesized understanding of information drawn from multiple information sources.

The lower scoring essay does not do these things as well. This participant identifies the topic at the start of her essay but does not specifically write with audience in mind, as requested in the prompt (i.e., a government considering nuclear energy implementation). She relied largely on one site to construct this essay ([nuclear-fission.blogspot.com](http://nuclear-fission.blogspot.com)) and listed facts, but struggled to integrate ideas across informational sources. She does integrate background knowledge, but there is insufficient evidence to conclude that she integrated information across multiple, multimodal texts.

*Example 2: Lower scoring essay [8/20] [Treatment participant, practice session 1]*

Nuclear fission is a type of nuclear power that provides energy for an alternate energy use. Nuclear power comes with many risks and disadvantages, although it does have positive effects as well.

Everything has both pros and cons, and nuclear power is no exception to this rule.

The pros of nuclear power include the low CO2 emissions, the amount of energy it provides (200 times that of burning coal) and its low operation expense.(1) Though its pros sound like good things the cons of nuclear energy far outweigh the good.

Nuclear power plants take up a lot of space and are very expensive to build, although once built they are inexpensive to maintain.(2) This type of power comes with many dangers, some of which are- the radiation waste is very hard to get rid of safely and is extremely harmful to both humans and the environment; the workers have a high-risk of a meltdown within the power plant that could lead to a nuclear explosion (3); the center could be the target of a terrorist attack, because of the damage a nuclear explosion can cause, and fission bombs (also known as atomic bombs) are a possible result from this process.(4)

The radiation waste from nuclear fission plants is extremely dangerous, and hard to store safely, this waste lasts for 200 to 500 years .(5) If an atomic bomb was created as a result of the power plant the effects could be devastating, in the bombing of Hiroshima, Japan, hundreds of thousands of civilians were killed as the result, 69% of buildings were destroyed, and 7% were damaged severely.(6)

Overall the effects of nuclear

### Annotations

1. From [nuclear-fission.blogspot.com](http://nuclear-fission.blogspot.com)
2. from [nuclear-fission.blogspot.com](http://nuclear-fission.blogspot.com)

3. all of the cons are also from [nuclear-fission.blogspot.com](http://nuclear-fission.blogspot.com)
4. all of the cons are also from [nuclear-fission.blogspot.com](http://nuclear-fission.blogspot.com)
5. also from [nuclear-fission.blogspot.com](http://nuclear-fission.blogspot.com)
6. From [wikipedia.org/Hiroshima](http://wikipedia.org/Hiroshima)

**Summary of section 2 results.** Analyses of students' use of trace indicators of integration in their written arguments revealed several important findings. At pretest and posttest, the groups' overall TII index scores did not differ. The treatment condition mean score did, however, change between pretest and practice session 1 when the intervention was introduced and the difference between pretest and practice session 2 was even statistically significant ( $Z = -2.384$ ,  $p = .017$ ,  $r = .84$ ). This result suggests that the LINKS intervention may have been a constructive influence for the treatment condition, enabling them to engage more indicators of integration once receiving the first step of the intervention. The large improvement in mean TII scores for the treatment group from pretest to practice session 1 was the most considerable change on this index for the treatment group (the small upward shift at practice session 2 made the mean difference statistically significant). For the remainder of the study, their TII index scores remained steady, never dropping back to their pre-test levels. Importantly, the control condition saw no analogous bump in TII at any point in the study. Reading, writing and simply practicing this type of reading-writing inquiry activity with a partner five times did not seem to boost these students' TII scores in a similar way.

Additional insights emerged from a close analysis of discrete criteria from the TII rubric. At pretest and posttest, both groups were equally likely to use more than one text to inform their written arguments. They were also equally likely to integrate corroborating facts gathered from two or more Internet texts. Where the groups differed, however, was in their use of counterpoints

gathered from Internet texts that were different from those used to inform the construction of their central written arguments. At pretest, the control group's mean rank (10.56) was statistically significantly higher than the treatment condition's mean rank (6.50) on this criterion ( $p = .042$ ). At posttest, the treatment condition's mean rank (10.50) was not strictly statistically significantly higher than the control condition's mean rank (6.44) ( $p = .053$ ) but the effect size ( $r = .48$ ) suggested a substantial influence of treatment. The within-group change for the treatment condition from pretest to posttest was, however, statistically significant. The effect size ( $r = .79$ ) suggested a very substantial effect of treatment for this group between pretest and posttest.

Leveraging the assumption that students' writing can be used to infer proof of multiple text integration processes, these findings suggest that the LINKS intervention did have a statistically significant impact on the treatment group's ability to pull contrasting perspectives gathered from a broader set of Internet texts into their written arguments. We can also infer that LINKS may have both cued students to talk about their background knowledge while reading, and enabled students to make more use of this information in their written arguments.

### **Section 3: Case Study Analysis of Dyad 8**

To further explore the impact of the LINKS intervention, particularly for the participants whose writing scores were lowest in this sample, I conducted a case analysis of Dyad 8, John and Alex (pseudonyms). These boys received the LINKS treatment over an eight-week period that began November 14, 2012 and ended January 8, 2013.

John, 14.6, and Alex, 15.0 years of age attended the public high school. Like all participants in this study, they were in the ninth grade. On their self-report surveys, John indicated a very high interest in science overall, rating his interests as 5/5 generally. Alex, however, ranked his interest in science at only 3 out of 5. Alex did report a stronger interest in



biology and earth sciences (4/5) than in chemistry and physics (1/5) however. Both boys reported having Internet access at home. John reported doing homework, reading sites of interest and social networking most often from home. Alex reported watching videos, doing research for school projects and other homework assignments most often. Both boys reported that creating multimedia presentations was the school-based online activity of highest frequency, followed by conducting Internet searches, doing projects and visiting teacher-recommended websites.

The boys' mean WRMT-III reading passages subtest score was 119, but John's score (104) was lower than Alex's score (134). Overall, John's score was among the lowest scores in the sample group and Alex's was among the highest. Interestingly, John reported that he liked to read, whereas Alex answered that he did not like to read in general. Both boys ranked their own reading abilities at 4/5.

**Evidence of change in writing.** At pretest, John and Alex's scores on the TII index were the lowest in the sample group. John's integration score was 1; Alex's score was 3. At posttest, however, both John and Alex's scores were higher, 7 and 10 respectively. At posttest, the boys drew on more information from a wider range of information sources to construct their persuasive arguments. At posttest, their arguments were more consistent with the topic prompt than at pretest. They also used more linguistic indicators of integration in their writing. Evaluations of their pretest and posttest essays are included in Appendix J. Here, however, is the markup of both boys' pretest and posttest essays with annotations provided below each essay in numbered lists. No edits have been made to their work.

***Pretest persuasive arguments.*** John and Alex wrote first on whether or not to accept the risks of radiation treatment.

*John's pretest written persuasive argument [1/20 on TII index].*

But if I were in the shoes of a cancer patient there are a few factors I would have to consider before making decisions. depending on my type of cancer, personally i would agree to the best treatment for myself. This is because I am a fighter for my friends and family, to spend more time with them.

Although many people have cancer, many do not have the same type or severity.(1)  
So side effects of the cancer and treatments are different for almost everyone (2), including the fact that all of our body chemistries are different. some factors that may apply to your specific side effects are age, your current/past health, your specific type of cancer, and your treatment plan(3).

### **Annotations**

(1) from patientresource.com

(2) a bit of an overgeneralization -- it's grounded in the notion, read at

[patientresource.com](http://patientresource.com) that side effects vary depending on age, health, specific cancer and treatment plan -- but I checked, and there was no mention of body chemistries and no mention that cancer and treatments are different for almost everyone at the site.

(3) taken directly from notes from [patientresource.com](http://patientresource.com)

*Alex's pretest written persuasive argument [3/20 on TII index].*

Do you think a person take risky procedures to cure or help cure cancer? I think this all depends on the person who has this cancer. for some people this would be a lifesaver and others it would be a painful agony till death. Why would this be a lifesaver or painful agony for some people and not others.

Cancer treatments have many side effects to only hope to cure cancer.(1) Not every person will have the same side effects from treatment depending on their age, health, type of cancer, and treatment plan.(2) Some possible side effects would be minor discomfort or inconveniences but there is also major side effects such as pain, major discomfort, or emotional distress. Doctors can give you advice on how to help these side effects(3).

there is certain tasks you can do to help prevent side effects of treatment.(4) you can take radioprotective drugs that help with the side effects of the treatment. It has side effects of its own so not all doctors recommend it.(5) these drugs can only be used if you have head or neck cancer,(6) but if you get plenty of rest, eat a balanced diet unless you have stomach or pelvic cancer. (7)taking care of affected zones of the skin with lotions or soap can help as long as you don't put heat or coldness to the affected zone.(8)

### Annotations

**General Note:** What I notice here is that he lists the facts from patientresource first and then the facts from cancer.org second – but he doesn't connect the ideas or make specific thematic links using information from both sources. Moreover, he doesn't construct an argument.

This is an integrative inference -- he did read about many side effects on the two sites they read closely.

(1) from [www.patientresource.com](http://www.patientresource.com)

(2) also from [patientresource.com](http://patientresource.com)

(3) I see this as a transitional phrase.

(4) from [cancer.org](http://cancer.org)

(5) from [cancer.org](http://cancer.org)

(6) These ideas come from [cancer.org](http://cancer.org) but strung together as they are, they make no sense.

(7) also from [cancer.org](http://cancer.org)

**Posttest persuasive arguments.** John and Alex wrote last on the issue of synthetic vs. natural soaps for babies.

*John's posttest written persuasive argument [7/20 on TII index].*

In the debate between natural and synthetic soaps there are a few variables that come into play. In my research i had found that the ingredients in some mass brands contain sodium laurel sulfate which causes eye damage in cases where the soap comes in contact with the eye.(1) On another side of the spectrum soap made with natural products wash off cleanly where synthetic soaps leave a residue.(2) However the choice between synthetic and natural soaps comes down to personal qualities; do you desire the need to have a soap that is in no doubt good for you (natural) or do judge your brand by price tag and the possibility for life changing outcome through injury?(3)

Personally i have found a decently cheap synthetic soap that works perfectly well every time. I suggest you do the same.(4)

### **Annotations**

(1) from dherbs.com

(2) from examiner.com/article/the-difference-between-natural-and-synthetic soaps

(3) This is an idea that they discussed after reading charearl.com/soap/naturalvssynthetic

(4) This is from his background knowledge.

*Alex's posttest persuasive argument [10/20 on TII index]*

For a parent trying to raise there child you have to worry about a lot of things but have you ever been concerned of soap. Mass produced soaps are generally synthetic and may cause damage to your skin and the eyes of infants and small childran.(1) Natural soaps are safe

for children and adults but they have their drawbacks too.(2) Synthetic soaps use man made chemicals like ultramarine and oxides in their soaps that are not found in natural soap.(3) Synthetic soaps take out glycerin in their manufacturing process that can make the soaps harder to wash off without it but natural is light because it has glycerin in it.(4) glycerin lets your skin breathe and release toxins.(5) Natural soaps cost more and can be relatively hard to find and may have to be ordered(6). Mass produced soap is easy to find and cheap.(7) I think Synthetic soaps are not as bad as they seem as a child I would sometimes get it in my eyes and it never did anything to my 20/20 vision. My whole life I used them and I have never had a side effect (8). Synthetic could be safer if you get a scrubber which can remove so off that left over soap on your infant's skin is removed.(9) I would suggest that you use synthetic soaps because of its availability and price but if your child does have problems try the natural.

### Annotations

(1) from [dherbs.com](http://dherbs.com)

(2) An integrative statement -- he read this at [dherbs.com](http://dherbs.com); [www.soap-making-resource.com](http://www.soap-making-resource.com); [examiner.com](http://examiner.com) and [charearl.com](http://charearl.com)

(3) from [www.soap-making-resource.com](http://www.soap-making-resource.com)

(4) from [charearl.com](http://charearl.com)

(5) also from [charearl.com](http://charearl.com)

(6) I think this is an integrated statement -- combining their impressions of the list of "more natural soaps" and the fact that they had never seen any of the natural soaps listed at [dherbs.com](http://dherbs.com) PLUS the information at [charearl.com](http://charearl.com) that identified the issue of commercial availability as a pro for synthetics.

(7) from [charearl.com](http://charearl.com)

(8) connection to bgk -- but not really informed by data.

(9) **Note:** He is focused on the one issue re: synthetics -- i.e., that it stays on skin and clogs pores -- but he doesn't really seem to have understood the key issue of infants having such delicate skin and that a scrubber could potentially harm a child as well. He definitely read about toxins in synthetic soaps -- and mentions ultramarine and oxide early in the essay -- but he has not integrated the idea of toxic chemicals in synthetic soaps and how these could harm a very sensitive/delicate baby. I wonder if this is a case of classic assimilation -- where he has taken the information that is consistent with his own pre-existing understanding and focused only on the facts that seem to affirm it.

John's final essay was more purposefully responsive to the prompt, engaged information read across three different texts and included the juxtaposition of different perspectives. Alex's posttest essay includes more integrative statements based on ideas extracted from more information sources. In his posttest, he acknowledges both sides of the issue and provides information from multiple texts according to this framing. In the pretest, he listed facts from one source and then facts from the other without integrating them.

Integration scores (See Appendix H for all evaluation rubrics and justifications for scoring of each criterion) show that both boys were able to construct a more integrated representation of their understanding from multiple Internet texts at posttest than at pretest, although, arguably, the boys would still be considered quite novice, or early in their learning trajectories at the end of the study.

**Evidence of change during online inquiry.** Given the observed improvements in writing, it is of interest to explore changes in the boys' application of strategies while reading as well. Although other factors may certainly have played a role, change in the boys' pretest vs. posttest

reading processes may have contributed to improvements in their TII index scores. Given that these boys demonstrated the least preparation for the writing task at pretest when compared to other participants in the study, teachers of students like John and Alex could benefit from understanding how, if at all, LINKS disrupted their patterns of strategy application during online inquiry.

One comparative metric of interest is the number of sites viewed at pretest and posttest. At pretest, the boys viewed four sites and read only two closely. At posttest, they viewed five sites and read four closely. Interestingly, they attended to twice as many texts closely, but only viewed one additional site at posttest. It doesn't therefore seem that the intervention lead to an increase in the total volume of texts that students accessed.

Consistent with the first research question, I also wondered to what extent LINKS may have shifted the boys' use of time as an indicator of change in strategies use. As I did for the larger groups, I calculated two types of time-focused data to explore the impact of LINKS on the boys' strategy use at each of the five sessions. First, I calculated relative duration for  $[(PST)^2 + (iC^3)]$  strategies, which is a measure of the proportional use of time (Bakeman & Quera, 2011, p. 98). I also calculated mean gap, which is the average amount of time between event onsets. This calculation is a proxy for differences in the speed with which the boys applied particular strategies over time and provides an indication of pattern shifts in reading process.

***Relative duration of strategies use.*** Figure 15 shows the relative duration of ten different codes at each of the five sessions in the study for Dyad 8. The similarity in the general shape of the graph is remarkable. Relative to one another, Dyad 8 applied similar proportions of their time to each of the strategies over the five sessions, suggesting that the task itself may have been the

most significant determinant of the processes they engaged and the duration of their relative engagement. There are, however, some interesting shifts in relative duration for particular codes.

Notice that for this dyad, the highest proportion of time (32 %) was spent identifying important information (G) at pretest, followed by constructing meaning within single texts (M) at 16 % and selecting sources (D) at 6 %. Updating understanding (J) accounted for 5 % of their time and Search (C) occupied just 4 % of their activity. At pretest, none of their time was spent on considering the Type (E) of text they were reading or on Connecting what they read to Other Texts (I). Just 1% of their time was spent on considerations of trustworthiness (F). Three percent of their time was spent comparing what they were reading to their background knowledge (H). Although this strategy is not represented in the graph, fully 25 % of the boys' time was spent taking notes at pretest. Though not plotted in Figure 15, none of their time was spent considering *Purpose (A)* or what they knew in advance of search (*Pre-Existing Knowledge, B*).

At posttest, there are obvious shifts. Less time (26%) was spent identifying important information in texts (G), although this strategy still accounted for the largest proportion of their time. Interestingly, a larger proportion of time (11 %) was spent on *Search (C)*, an increase of 7 % from pretest. Compared to pretest, they also doubled the proportion of their time spent selecting sources (*D*) (12 %) but spent 5 % less time (11%) constructing meaning within single texts (*M*). All of these trends are consistent with those observed in the larger group.

For Dyad 8, Strategies *Compare to background knowledge (H)*, *Connect to other texts (I)* and *Continually Update (J)* all saw a proportionate increase. This is a departure from the trends seen in the larger group, which decreased relative duration on *Compare (H)* and *Continually Update (J)* and increased only slightly on *Connect (I)*. Taken together these three strategies accounted for 21 % of the boys' time at posttest, compared with just 8 % at pretest. At pretest,



the boys spent no time connecting to other texts; at posttest, they did this with 2 % of their time. At pretest, the boys spent just 3 % of their time making comparisons between the texts they were reading and their background knowledge. At posttest, this increased to 8 % of their time. Most significantly, the boys spent 11 % of their time updating their understanding at posttest, an increase of 6 %. Also remarkably, at posttest, just 9 % of the boys' time was spent taking notes and 5% of their time was spent considering *Purpose (A)* and *Pre-existing Knowledge (B)* before they engaged in search activities.

It's also interesting to consider that the boys spent no time at posttest considering the *Type* of texts they were reading (*E*) or evaluating *trustworthiness (F)* explicitly. During practice sessions, when I was scaffolding their reading with guided questions and prompting them to focus on the  $[(PST)^2 + (iC^3)]$  strategies, they did engage in these activities. However, at posttest, in the absence of instructional scaffolds, they did not. This finding is consistent with analyses of relative frequencies between and within groups. Overall, treatment participants did engage relatively more *Type (E)* and *Trustworthiness (F)* strategies than the control participants (See Figures 4 and 5) during practice sessions, but differences in relative frequencies were not statistically significant at posttest. Generally, the treatment participants were more likely to engage these critical evaluation skills when prompted through the LINKS intervention than at posttest when the intervention was removed. Interestingly, for the larger group, the Friedman's ANOVA showed a statistically significant increase in *Trustworthiness (F)* episodes within the treatment group over time. John and Alex, however, spent no time engaged in trustworthiness activities at posttest. It may be that critically evaluating information sources is especially challenging for students like John and Alex who are rather novice, or at very least, present as very early in the developmental learning trajectory for these skills.

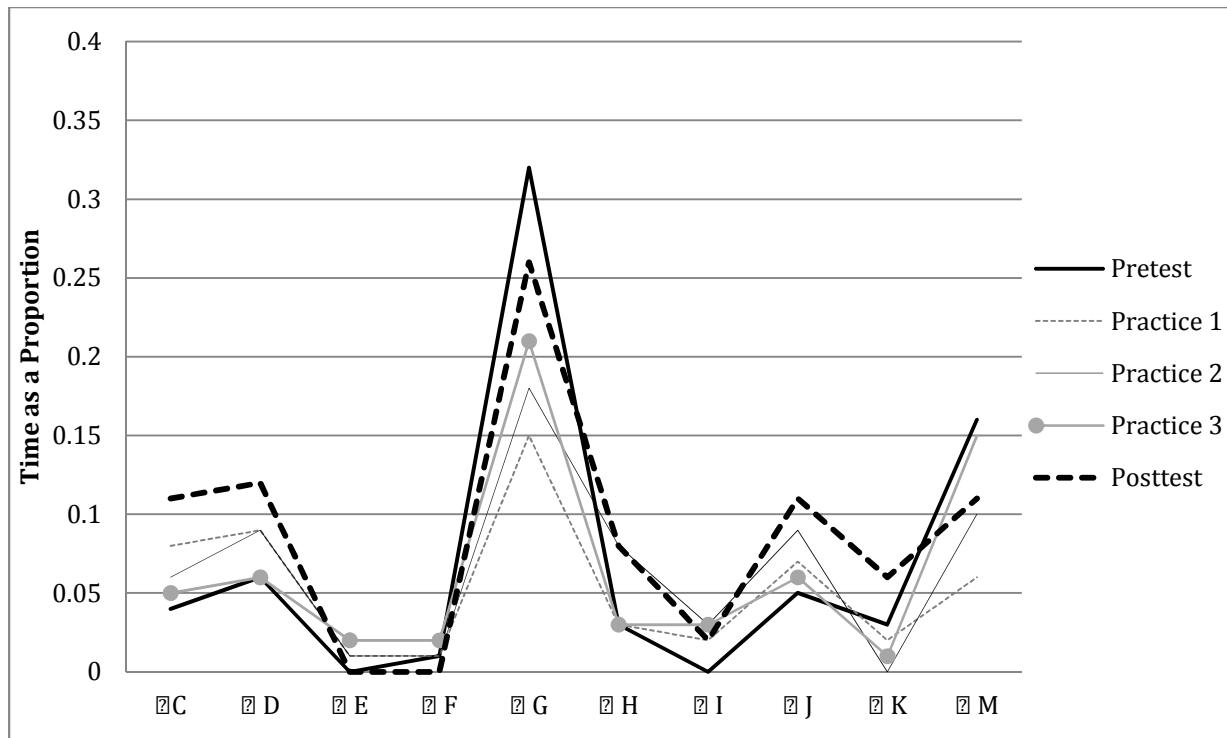


Figure 15. Relative duration of strategy application for Dyad 8 over time.

Note: C = Search, D = Source Selection, E = Type, F = Trustworthiness, G = Identify Important Information, H = Compare to background knowledge, I = Connect to Other Texts, J = Continually Update, K = Student Questions, M = Construct Understanding within a Single Text. Times are expressed as proportions of 1, or percentages.

**Mean gap of strategies use.** Change in the mean time between events of each type suggests a shift in pattern of activity. Figure 16 shows differences in Dyad 8's mean Gap time for thirteen strategic activities. As in figure 15, pretest and posttest lines are most prominent to highlight these differences. The graph shows considerable variability across practice sessions. However, from pretest to posttest, mean gap times decreased for four strategies: *Search (C)*, *Source Selection (D)*, *Compare to Background Knowledge (H)* and *Continually Update (J)*. This suggests that Dyad 8 initiated use of these strategies in quicker succession at posttest than at pretest. Conversely, the mean gap time for *Student Question (K)*, *Construct Meaning from Single Text (M)* and *Notetaking (N)* increased from pretest to posttest, suggesting that the mean time between onsets of these strategies lengthened. Taken together, it seems that at posttest, Dyad 8

did activate, in quicker succession, the strategies that would predict construction of an integrated understanding (i.e., C, D, H and J) vs. those that would predict more protracted focus on a single text (i.e., M and N) or questions about how to do things (K).

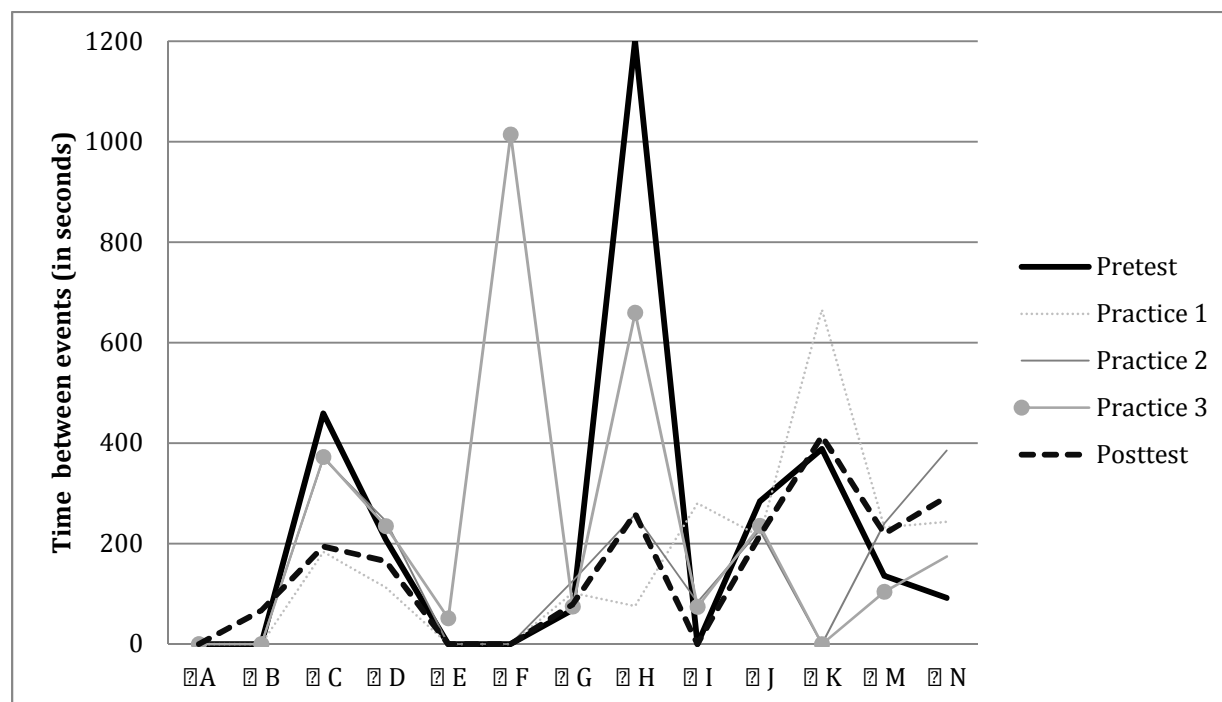


Figure 16. Mean gap for Dyad 8 at each point in the study.

**Qualitative comparison of integration in Dyad 8's thinking.** I compared pretest and posttest transcripts and my annotations of their reading processes to explore the qualitative differences in the ways these boys were constructing meaning from multiple texts. At pretest, I was struck by the school-based script these boys seemed to be following. After finding their first text to read closely (the cancer.org website), and before taking notes from it, John asked, “Do we cite our sources?” My response was non-committal. “You can, if that feels good, then do that. It’s your choice [shrug].”

In their notes, the boys wrote American Cancer Society and Cancer.org/American Cancer Society at the top of their page, but did not record the name or URL for the second site they read closely. In my notes, I reflected on this question and what I felt it meant.

*The question, “do we cite our sources” is never, it seems, because the kids want to engage in a critical or evaluative conversation about the author’s perspectives and whether or not they are trustworthy. Rather, it is my impression that this question is largely grounded in a school-based “script” for such inquiry-based experiences. They ask because they’ve been told they should do it -- but not because they’ve developed the evaluative dispositions around inquiry that they need to develop. Instead, they just want to do the “right” or the “expected” thing.*

Interestingly, even though John & Alex asked about citing sources, and wrote down “American Cancer Society cancer.org” in their notes at pretest, they did not explicitly engage strategic processes indicative of attribution, critical evaluation of authorship, or type of text. This was the only time they mentioned anything connected to the evaluation of a text’s *trustworthiness (F)* but the activities that followed suggested they had not internalized important mindsets or dispositions around critical evaluation.

I also noted how the boys engaged in a patterned process of note taking that I called, “find information, take notes, find information, take notes” which seemed to mimic the scribing of notes from a textbook. As average and above average readers respectively (based on their WRMT scores), John and Alex showed that they were good at summarizing, paraphrasing and constructing meaning within texts. However, they did not actively connect ideas across texts (or show much evidence of integration in their essays). In fact, they read only two texts during this session.

Here's what I wrote:

*As I'm coding identify important information (G) and student takes notes (N) - I find I'm back and forth again with these kids because they take notes and then talk about what they're reading. These guys are talking aloud which means they show me how their processes are switching. Note taking -- then finding more information -- then note taking -- then summarizing -- then questioning.*

Here's a representative excerpt that demonstrates the pattern:

G 06:34.2 [[www.cancer.org/treatment/treatmentsandsideeffects/physicalsideeffects/radiationtherapyeffects/index](http://www.cancer.org/treatment/treatmentsandsideeffects/physicalsideeffects/radiationtherapyeffects/index)]

ALEX: Alright. Let's read it, before we cite it. [suggesting that he wants to know what the text is about before deciding to use the information?] [reading silently]

6:52 ALEX:[reading text aloud] Preventing. I guess that could be good. [moving the cursor over the text]

JOHN: So, preventable and manageable common side effects [an interpretation of the title which is Preventing and Managing Common Side Effects]

7:13 JOHN: What is this thing

ALEX: I was writing preventable side effects

JOHN: What is this thing, though. American Cancer Society?

ALEX: [www.cancer.org](http://www.cancer.org)

7:38 JOHN: Ya. American Cancer Society

[they're not really questioning the site's trustworthiness, so far as I can tell from their interactions. Rather, HU is just wondering what the site is so that he can write down a citation. For this reason, I've just coded this as identify important information. He recognizes value in knowing what the site is -- but he doesn't

- question it or comment on it as a trusted source of information]
- D 06:56.9 [They click on the hyperlink -- which then takes them to a sub-page focused specifically on managing and preventing side effects of radiation treatment]
- N 07:29.1 Both boys put their head down and start writing the site URL ALEX: [writes] American Cancer Society JOHN: [writes] cancer.org/American Cancer Society [the sub-page loads]
- G 07:36.0 [They're now at the SP: /radiation/understandingradiationtherapyaguidefor patientsandfamilies/]
- 7:56 JOHN: Alright.
- ALEX: Alright. Preventable
- [They look at the text and take notes directly from it]
- 8:10 ALEX: How do you spell society?
- JOHN: s-o-c-i-e-t-y
- 8:28 ALEX: Alright. Radioprotective drugs.
- K 08:10.6 ALEX: How do you spell society?
- JOHN: S-o-c-i-e-t-y
- G 08:27.1 ALEX: Alright. Radioprotective drugs.
- JOHN: [reading from text directly and aloud] [ii1] When radiation damages nearby healthy tissue, it causes side effects. Many people worry about this part of their cancer treatment. Before treatment, talk to your doctor.
- ALEX: The only way to reduce side effects is by using
- JOHN: The radioprotective drugs
- N 08:55.6 ALEX: Right, so, radioprotective drugs. [he writes this down]
- G 09:17.8 JOHN: [reading aloud from text] [ii1] These are drugs that can be given before radiation treatment to protect certain normal tissues in the treatment area. The

one most commonly used today is amifostine.

M 09:35.9 ALEX: So, right. It's taken before radiation treatment?

JOHN: Hmhm.

At posttest, these boys did not engage *Type (E)* or *Trustworthiness (F)* strategies at all, but they did seem to cycle through a broader set of strategies that enabled them to construct meaning both within texts (which they did at pretest very well) and among them. My first comment was about how metacognitive these boys were as they read. Their processes engaged me in an important methodological reflection as well.

Here's what I wrote:

*They are REALLY metacognitive.*

*Methodological question: With these two I'm finding that they're generally so talkative and metacognitive that even though they're at the same website, I feel like they're switching between cognitive processes. Between 10:12 and 15:22 they were on the same site but I coded "identify important information" four times because they would identify an important fact, take notes, make a connection to background knowledge about it or summarize their understanding and THEN go back to finding more information. So, I felt like each time, there was a break in process due to another process taking over. Yes -- they take notes at the same site for this whole time, but unlike Dyad 4, for instance, where they would just ask "can I scroll down" or "are you done with this section" as they read silently, these kids are really very active in construction of meaning together. So, I really did feel that there was active switching of activities that justified the use of the same code repeatedly while at the same site.*

This next representative excerpt shows a broader range of strategic activities over the same amount of time (approximately three minutes) that starts with *Identify Important Information (G)* as well and includes a connection to background knowledge suggestive of analogical reasoning:

G 22:09.2 [ @ charearl.com/soap/Natural-vs-Synthetic-Soaps.html]

ALEX: And, they're easier to come by, too. [Which, I think I can say he understood from the snippet text for this page -- and he has integrated this in to his understanding of the pros of synthetics]

[scrolls down the site a bit]

JOHN: Ya. [silent reading]

M 22:41.9 JOHN: It says that commercial soaps lack in glycerin. It's removed because it decreases the shelf life of the product.

K 22:45.4 ALEX: Where do you see that?

JOHN: It's all in here. [question focused simply on where to find the information that HU identified as important]

J 23:16.3 ALEX: Well, that's a pro and a con [in response to HU's reading that glycerin is removed to prolong the shelf-life of commercial soaps] cause it make it, so you you have it longer but. It's the twinkie [an analogy to the shelf-life of the twinkie and all of the chemicals that allow them to last for such a long time] ALEX: Twinkies are so good, but they're going away.

JOHN: Huh? They're not going away are they?

ALEX: Yup.

JOHN: Why?

ALEX: Hostess went out of business. But. Back to this.

JOHN: They, they got bailed out.



- H 23:26.2 [they resume reading silently on the website]
- 23:53 JOHN: Oh well. They last for 26 years. {I coded this this way because they used something they knew about -- twinkies -- to understand shelf-life in soap}
- G 24:20.6 [they resume reading silently]
- JOHN: Commercial soaps also get rid of natural oils on our bodies as well that instinctively protect our skin from harmful elements and deterrents.
- ALEX: And they do not, er, ah, natural doesn't cover up the pores, allowing it to ah, excrete toxic waste. So, oh. We need a new pen.
- N 25:02.5 [They both take notes from this site, [charearl.com/soap/Natural-vs-Synthetic-soaps.html](http://charearl.com/soap/Natural-vs-Synthetic-soaps.html)]
- ALEX: [writes] Synthetic is cheaper and more convenient but natural soap is safer. Natural soap is lighter letting your skin breath but synthetic can cover up skin [sic]
- JOHN: [writes] Synthetic = cheaper + more convenient natural soaps = Big \$/ better for skin Natural qualities
- M 25:11.2 ALEX: So, first of all, I'm just gonna write what's down here. Synthetic is a cheaper more convenient but then if you want to, er if values don't bother you, go for the natural [Connection to other text? I?]

There are moments in this excerpt that qualify the kinds of integrative thinking these boys were doing as they constructed an understanding of texts. There is a broader range of codes in this excerpt, suggesting more diverse activation of  $[(PST)^2 + (iC^3)]$  strategies, but also construction of knowledge by seeing “pros and cons”, leveraging their understanding of

Twinkies as a way to understand glycerine as a preservative in soap, and at the end, making a statement that suggests they were connecting back to an earlier idea about who might object to using soaps containing animal fats (Alex suggested vegans wouldn't like that). Clearly, there is more going on here than just the reading and construction of understanding from individual texts mostly seen at pretest. Their move toward analogical reasoning and explicit connections to background knowledge definitely suggests more active construction of an integrated understanding of this topic.

I also noticed differences in their ability to use purpose to drive their process. At pretest, Alex tried to prompt a focus on purpose. He said, "I think it's good enough. I mean, we have to go to other stuff too." And then a minute later, after considering the Search Engine Results Page, said "We have to know what the actual side effects are, though." They ended up at the [www.patientresource](http://www.patientresource.com) site, but here's what I noted.

*'What I'm finding really interesting in this one is how the kids really lost their way. They found a website [www.patientresource.com](http://www.patientresource.com) that told them some general information about short-term, long-term and late side effects of radiation, but they never actually find/focus on information about the actual side effects. Alex wanted to do this -- but somehow, they lost their focus. It's like they got on this train and they didn't stop to question where this route was going to take them.'*

In fact, the boys concentrated on generalities, recording facts such as "side effects depend on your age, your overall health, your specific cancer, your treatment plan" and "some cause minor inconveniences or discomfort, or more discomfort, pain and emotional distress" but they did not record specific side effects of radiation treatment.

At posttest, they were much more deliberate to update and maintain their focus on purpose. In the excerpt above, Alex says “Back to this” when John, started bemoaning the demise of the Twinkie. At 20:44, Alex updates his understanding, with expressed focus on the purpose of their task that leads to some integrative thinking, and to the generation of a new search term.

J 20:44.7 [in response to the content that comes after the info re: fragrances]

Well, we don't need to know the process of making soap, so I think that's all we're going to get out of this website. [he clicks back to Google list of search results] We still haven't found the pros of synthetic soaps.

JOHN: The only pros that I think would be to synthetic soaps is mass production and [...] cheaper.

ALEX: cheaper.

C 21:08.2 ALEX: [types in new ST: is synthetic soap cheaper]

This idea of cost appears in both of the boys' posttest persuasive essays. At posttest, it seems the boys were able to maintain better focus on the questions that they needed answer through their inquiry activity. Moreover, they were also able to identify topic relevance in texts so that they could extract information that answered their questions.

**Summary of section 3 results.** I examined the particular case of Dyad 8 because John and Alex were the participants who, based on pretest TII scores, appeared to be earliest in their learning trajectory for multiple Internet text integration skills development, and therefore had the most to potentially gain from the LINKS intervention. The boys' written arguments did improve from pretest to posttest on the measures of integration itemized in the Trace Indicators of Integration (TII) rubric.

The analysis of relative duration and mean gap for  $[(PST)^2 + (iC^3)]$  from pretest to posttest reveals shifts in their processing patterns that may have contributed to the improvements seen in their TII Index scores at posttest. In particular, the pair spent proportionally more time engaged in *Search (C)*, *Source Selection (D)*, *Compare to Background Knowledge (H)*, *Connect to other texts (I)* and *Continually Update (J)* strategies at posttest than at pretest. The boys spent much less time taking notes at posttest than at pretest and they shortened the mean gap times between onset of *Search (C)*, *Source Selection (D)*, *Compare to Background Knowledge (H)* and *Continually Update (J)* strategies too, suggesting a quicker cycle of engagement with these integration strategies.

Qualitatively, the boys' conversations about their online reading and multiple Internet text integration processes changed. Compared to their pretest inquiry activities, at posttest they demonstrated a heightened ability to remain focused on task, to engage a wider range of strategic activities, and to leverage analogical reasoning, connections to other texts, and connections to background knowledge.

Although generalization is not possible from a single case, these results suggest that for students who share Dyad 8's profile, LINKS may alter or disrupt pre-existing patterns of multiple Internet text integration in ways that enable more evidence of integrative processing to show up in their written arguments.

Of particular note is the fact that relative duration and mean gap times of some processes changed whereas others did not. In particular, John and Alex did not explicitly evaluate text *Type (E)* or *Trustworthiness (F)* at posttest. For learners like John and Alex, more time, practice, think aloud modeling, and guided questioning with scaffolds may be required to enable them to engage critical evaluation strategies such as *Type (E)* and *Trustworthiness (F)* while reading online

without teacher support. This is important because the larger group did seem to engage these strategies more frequently than John and Alex did at posttest.

Another question that emerges from this case is why, given the changes observed in the boys' online inquiry processes, their posttest essays didn't show even more evidence of integration. For instance, the boys did not engage with multimodal texts at posttest, and their conclusions were rather more focused on opinion than on a careful examination of the evidence. General writing skills may have played a role here. Certainly, the boys might have benefitted from an even longer and more extensive intervention focused on developing these aspects of their persuasive writing.

## CHAPTER 5

### General Discussion

This is a study about an intervention, designed to support students' development of multiple, multi-modal Internet text integration skills. The impact of the intervention was examined in three ways: (a) by measuring frequency, relative frequency, relative duration and mean gap of students' application of strategic processes, (b) by investigating trace evidence of integrative processing in students' written arguments, and (c) by examining the processes and written arguments of one dyad who demonstrated little evidence of integration in their written arguments at pretest, but considerable improvement at posttest.

Based on results, the general theme of the LINKS intervention story is one of “disruptive promise”. I use the term disruptive promise to portray the way that LINKS nudged participants toward new patterns of integrative action during online inquiry for the purpose of writing a persuasive argument. All students started the study by applying a remarkably consistent set of strategies in remarkably similar ways, but as the intervention progressed for the treatment condition, indicators signaled that the consistency of these initial strategies was disrupted. What were these signals? Shifts in relative strategies application, change at certain moments in the study, and evidence of statistically significant change in the use of background knowledge and contrasting evidence in written arguments signal that the intervention holds promise as a method for teaching multiple, multimodal Internet text integration and that it warrants future investigation in classrooms with more participants and teachers. Given the complexity and the importance of multiple, multimodal Internet text integration skills development, I argue that even a “nudge” toward increased integration is an important finding. It is an especially important finding for adolescents on the precipice of significant cognitive and neurological change (e.g.,

Kuhn, 2006; Steinberg, 2005) who need to develop these skills during this critical developmental period to become more fully literate in the digital age. It is also especially important to recognize promise at a time when the curricular expectations for digital literacies instruction have moved more quickly than the generation of the evidentiary substrate on which digital literacies pedagogies should be built. Moreover, given findings that many adolescents struggle to construct an integrated mental model of understanding from multiple texts (e.g., Cerdán & Vidal-Abarca, 2008; Sevensma, 2013; Wineburg, 1991) this study offers teachers a promising point of departure for current and future instructional practice.

Recently, Colwell, Hunt-Barron & Reinking (2013) argued that pedagogies which develop “ingrained, spontaneous use of strategies for locating and evaluating information on the Internet when completing academic tasks” (p. 314) are especially challenging to cultivate, and that “spontaneous transfer to more authentic tasks is the acid test that should be the measure of an intervention’s success” (p. 315). The evidence analyzed in this study aligns with these assertions, but offers LINKS as a method well suited for further investigation. Evidence presented in this report shows that the LINKS intervention did enable the ninth graders in this study to engage some strategies relatively more often than control group participants, and in the absence of instructional support. In what follows, I address the research questions by summarizing evidence of the disruptive promise reported in the results.

### **Research Question 1: Impact on Strategy Application**

LINKS had no statistically significant impact on the total number of strategies that participants engaged at any point in the study, both within and between groups. On average, participants used the same number of strategies at each moment in the study, although descriptively, mean values at Practice Session 2 and Practice Session 3 (i.e., when the

intervention was administered) were more than 10 points higher for the treatment group than the control group, suggesting slightly increased strategic activity overall. This is the first hint of disruptive promise.

Despite the different instructional experiences during practice/intervention sessions for both groups, pretest and posttest comparisons of strategy-by-strategy codes revealed a common structure in strategy application by frequency. The three most frequently applied strategies at posttest for both groups were *G: Identify Important Information*, *D: Source Selection* and *C: Search*, in that order.

The structure of frequencies for *Identify Important Information (G)*, *Source Selection (D)*, and *Search (C)* are consistent with the fundamental processes that Rouet's (2006) Task-Based Relevance Assessment and Content Extraction (TRACE) model of search (p. 105) describes. Interestingly, both groups applied these strategies, even though the control condition was not explicitly instructed to do so, as the treatment dyads were. To construct a Documents Model of understanding, Rouet (p. 105) asserts that online readers must (a) construct their model of understanding based on internal needs and environmental constraints, (b) assess document relevance based on available information resources and search tools [i.e., *Search (C)* and *Source Selection (D)*] and (c) extract and integrate content information recursively [i.e., *Identify Important Information (G)*].

Given that the LINKS intervention included the modeling of these three strategic processes through think alouds, guided questioning and the general structure of the [(PST)<sup>2</sup> + (iC<sup>3</sup>)] framework itself, but both groups engaged the same three strategies most, second most and third most, the task itself may have driven the structure of students' responses to it. I discuss this further below. Interestingly, however, the treatment group increased its application of *Source*



*Selection (D)* activities by a larger margin than the control group and decreased its use of *Identify Important Information (G)* by a larger margin as well. So, although participants were all engaged in the same fundamental TRACE processes described by Rouet, LINKS may have moved treatment participants toward an increased focus on *Source Selection (D)*, a necessary precursor to the evaluation and reading of multiple texts vs. the control condition which did not demonstrate the same degree of shift. This is a second indicator of disruptive promise.

Importantly, there is an empirical basis in the online reading comprehension literature for the “task” hypothesis of patterns of strategies application. Zhang & Duke (2008) found consistencies among expert online readers on the sequence of strategies applied for each of the tasks they assigned (i.e., searching for specific information vs. searching for general information vs. reading for entertainment). These researchers also found that expert online readers altered strategy application according to reading purpose. Although certainly less expert than the readers in Zhang and Duke’s study, the ninth graders in my study may have also used the task to frame the structure of their reading activities in ways that enabled them to accomplish their reading-writing purpose. John and Alex (Dyad 8) were better able to focus on their task at posttest as well and, perhaps not surprisingly, demonstrated more trace evidence of integration in their posttest essays. These findings beg many questions about the interactions of task and strategies development for the construction of meaning from multiple texts online. In particular, I wonder if the modeling of task focus, or even simply prompting students to revisit their task purpose more often would be an effective intervention for supporting synthesis in and of itself.

The third indication that LINKS had a disruptive effect on pretest patterns of strategies application during online inquiry emerged from between-groups comparisons of relative frequencies by strategy. Treatment participants applied *Type (E)* and *Trustworthiness (F)*

strategies relatively more frequently during Practice Sessions 2 and 3 than control participants. According to Rouet (2006) source information gives readers a framework for comparing content and permits relative weighting when information from many sources differs (p. 74). I predicted that an instructional focus on content, source and the relationships among documents could move novice online readers toward a more expert level of multiple text integration. It is unclear from the analyses conducted whether students *weighted* certain evidence more heavily, as Rouet suggests occurs, but it is clear from the results that treatment participants were more focused on critical source evaluation strategies when they were receiving LINKS than when they were not. Moreover, they demonstrated more trace evidence of integration in their essays during practice sessions 1 and session 2, when LINKS included the most instructor scaffolding, than they had at pretest.

This finding raises several questions. Are these critical evaluation strategies more important to multiple, multi-modal Internet text integration than others? Or, are they tractable for most ninth-graders only when support is given? It may have been the case that the participants had been taught to critically evaluate texts during online inquiry in school (a few students mentioned never using Wikipedia, and not to trust blogs, for instance; John did ask if he should “cite sources”) so that when they were reminded to think about text type and trustworthiness during the study, it was easier for them to engage these strategies over others. Future research should address questions about the relative contributions of *Type (E)* and *Trustworthiness (F)* strategies to integration in written arguments. Such studies would help to clarify, for teachers, which strategies are most essential to teach in service of multiple, multimodal Internet text integration skills development.

## **Research Question 2: Impact on Trace Evidence of Integration in Persuasive Arguments**

A fourth indication that LINKS holds disruptive promise for instruction of multiple, multimodal Internet text integration skills emerged from analyses of trace indicators of integration in participants' essays. No statistically significant between groups differences were found at any point in the study using non-parametric tests on the overall TII index. Results of Wilcoxon signed-rank tests showed, however, that there was a general and statistically significant trend toward integration in the treatment condition that was not concomitantly observed in the control group. This fourth piece of evidence is interesting and important because, as noted, LINKS is not a "writing" intervention per se.

Most importantly, when scores on discrete criteria from the TII index were analyzed, the fifth and strongest indicator of LINKS' promise as a disruptor of pre-existing multiple text integration habits emerged. Although both groups were equally likely to use facts from more than one text, and to cite corroborating facts in their persuasive arguments, the treatment group was more likely to integrate their background knowledge in their essays. The treatment group was also more likely to integrate counter points to the main argument in their essays that were gathered from texts not used to construct the main argument. This evidence, in particular, is especially suggestive of LINKS' impact on processes that enable the construction of an integrated mental model of understanding from a broader set of texts representing a broader range of perspectives.

The case of John and Alex shows change in online inquiry processes from pretest to posttest and an increase in trace indices of integration in their essays too. Moreover, by posttest, these boys no longer copied notes as though from a textbook, but demonstrated a broader range of integrative thinking strategies, including analogical reasoning, made explicit connections to

background knowledge, and used pro-con heuristics to integrate understanding of multiple texts. This is the sixth indicator of disruptive promise.

Though generalizable claims are not possible here, these six hints of change and movement toward integration for the treatment condition, in ways not observed in the control, justify the claim of promise and future investigation in a broader set of learning contexts.

### **Why was LINKS disruptive?**

The foundation for the claim that LINKS disrupted pre-existing patterns of integration and inquiry has been built from multiple pieces of evidence. This method of evidentiary bricolage builds a foundation for future study of LINKS as a method that could be tested in a range of classroom contexts and across a range of content-area disciplines (not just science). The question that arises, however, is why LINKS might have had any effect at all? What mechanisms might have been at play to enable treatment participants to engage different sets of strategies and to demonstrate more trace evidence of integration in their writing? I present four connections to research worthy of future study.

Returning to Rouet's TRACE model, a strong theoretical influence for the design of LINKS and the  $[(PST)^2 + (iC^3)]$  strategies, the LINKS treatment may have influenced students' understanding of their "internal needs and environmental constraints" (Rouet, 2006, p. 105) in ways that enabled them to (a) engage more *Type (E)* and *Trustworthiness (F)* strategies during practice, (b) to apply more trace indicators of integration in their persuasive essays once introduced to the intervention, (c) to engage more explicit evidence of background knowledge during inquiry and in their persuasive arguments at posttest, and (d) to use more counter facts at posttest than the control group. Screencast think alouds and guided questioning may have provided referents against which "internal needs and environmental constraints" could be judged.

Likewise, the environmental constraints of the process may have been more obvious to treatment dyads because these were modeled and scaffolded more clearly for them.

Alternatively, and this is the second connection to research, LINKS may have enabled treatment participants to construct meaning in a task that was generally very open, and on topics they knew relatively little about in ways that the control experience did not. McNamara and Shapiro (2005) found that the construction of a cohesive situation model from multiple linked hypertexts was dependent on the structure of the hypertext environment itself, but also on the reader's pre-existing domain knowledge. Readers with more content knowledge in McNamara and Shapiro's study were more able to construct meaning in open hypertext systems whereas readers with less content knowledge benefitted from hypertext environments that explicitly cued the relationships among texts. For treatment participants in this study, the strategic cueing of background knowledge taught during LNKS, along with the strategies for search, critical evaluation, comparing, connecting and continually updating may have been enough to enable use of whatever background knowledge students had during the reading and writing phases of the inquiry task so that they could construct a more integrated understanding of the topics. Although posttest Trace Indicators of Integration (TII) index scores did not differ between groups, treatment participants' scores *did* improve by a statistically significant margin by the second practice session and remain at that level for the remainder of the study. Treatment participants also used more background knowledge in their essays. In brief, the strategic content (i.e., the what) of the LINKS intervention may have equipped these students to leverage even limited *a priori* domain knowledge in an open online inquiry activity.

Thirdly, we might invoke theories of cognitive load (e.g., Kirschner, Sweller & Clark, 2006; Mayer, 1979, 2004) and self-regulation in hypermedia learning (e.g., Azevedo &

Witherspoon, 2009; Negretti, 2012). It may be that the modeling and guided questions provided to treatment participants freed their working memories so that during writing, they could leverage their background knowledge as a point of comparison for the information they had gathered, and integrate counterpoints from a broader range of texts. Certainly, the intervention was designed to scaffold precisely the skills that Azevedo & Witherspoon (2009) identify as essential for self-regulated learning, understanding, and problem solving in hypermedia contexts, namely, “planning processes such as activating prior knowledge, setting and coordinating sub-goals that pertain to accessing new information [...] coordinating several informational sources, generating hypotheses, extracting relevant information from the resources, re-reading, making inferences, summarizing, and re-representing the topic based on one’s emerging understanding through taking notes and drawing” (p. 321). If we substitute this last idea— taking notes and drawing—with writing, then I contend that LINKS, with its protocols parallel to those outlined by Azevedo & Witherspoon may have scaffolded self-regulatory processes for treatment condition participants in ways that supported greater integrative thinking at certain moments during the study, including at posttest for two key criteria of integration. This hypothesis is speculative, of course, but it resonates because statistically significant between groups differences in relative frequencies application occurred during practice, at moments when the intervention was scaffolding treatment participants through the complex process of multiple, multi-modal Internet text integration. At these moments, for a select few strategies it seems that LINKS shook up the status quo and enabled more critical evaluation of texts. Even though posttest differences on frequency, relative frequency, relative duration and mean gap did not differ statistically for most types of strategic episodes between groups, treatment participants did engage background knowledge at posttest more often than control participants. Moreover, a large

and statistically significant effect was found for use of background knowledge and counterpoints in the treatment group's persuasive essays at posttest, suggesting that practice *with LINKS* (vs. the control experience) may have lessened the cognitive demands of the task so that treatment students could bring together more pieces of information from their own experience and from their inquiry activities than they were able to use at pretest.

Fourthly, the interaction of writing and the LINKS intervention during online inquiry definitely warrants further investigation. Consistent with research on writing-to-learn in school (e.g., Klein, 1999; Klein & Rose, 2010; Langer, 1986a, 1986b; Newell, 2006) I hypothesized that the writing of persuasive arguments in this study might support multiple text integration, but I didn't teach writing, nor did I focus too closely on note taking processes in my analyses, other than to record that students were taking notes as part of their inquiry processes. For the control condition, there is insufficient evidence to conclude that the act of writing and/or note taking was, itself, a sufficient scaffold for integration. Although they started off the study able to demonstrate slightly higher mean levels of integration in their essays than the treatment condition, the control group showed no change in integration over the five sessions of the study. For the treatment condition, however, the LINKS intervention seems to have enabled them to bring together more information, from more texts, including those that presented contrasting viewpoints, and apply more linguistic markers of integration such as transitional phrases and parallel structures after each of the first two practice sessions and at posttest. Having received guided support and think aloud modeling of multiple text integration via screencast, the act of writing may have enabled integration of multiple, multimodal Internet texts for the treatment condition in ways that were inaccessible to the control condition who received none of the integration supports during online inquiry.

It is also important to consider the potential of LINKS for students who seem to be very early in their learning trajectory for online inquiry and multiple text integration. Dyad 8, the treatment condition students who, on the basis of their pretest trace indicators of integration (TII) writing scores presented as most novice, also improved most notably over the course of the study. John and Alex's TII scores went up six and seven points respectively between pretest and posttest, the largest range measured in the study. At posttest, they spent relatively more time on strategies that would predict multiple text integration—specifically, *Continually Update (J)*, *Connect to other texts (I)* and *Compare to background knowledge (H)*. Like the larger group, they also spent more time engaged in *Search (C)* and *Source Selection (D)* activities. At the end of the study, these students would still be considered “novice” but the LINKS intervention may have hastened their development along a learning trajectory. Future research should very deliberately focus on students who show the least preparation for multiple, multi-modal Internet text integration at pretest in order to more fully understand the potential of LINKS to scaffold their growth. Arguably, these are the types of students for whom intervention is most essential. Given the increased demands for integration in schools outlined by the Common Core State Standards in the US (National Governor’s Association Center for Best Practices & Council of Chief School Officers, 2010), students like John and Alex who receive little support for online inquiry, as in the control condition, may lose out on important chances for online literacies development in high school that would prepare them for the advanced literacy demands of college, career, and citizenship in the twenty-first century.

**Questions requiring further analysis.** Returning to the theoretical foundations for this work, I cannot confidently claim that treatment participants engaged in more *rapid* criss-crossing of the web landscape in ways that Spiro & DeSchryver have described (2009) to construct an



integrated understanding of multiple texts, although treatment participants did come to engage Search and Source Selection activities more often, constructed their persuasive arguments using facts from a broader set of websites, and as outlined in the description of Dyad 8, John and Alex did engage more and shorter cycles of strategy activation that included search at posttest. There were shifts in these skills, but more analyses will be required to thoroughly explore patterns of criss-crossing as generative.

Although I began with a plan to document “schemas of the moment” (p. 116) I found it difficult to know when a comment might be generative vs. a connection to background knowledge (code H). As an example, there was a moment at posttest when Alex, upon reading about animal fats in synthetic soaps, said “Well, I mean, that’s only if you’re vegan. It doesn’t really harm you, it’s just that they don’t like to use.” I wondered whether this was a generative schema of the moment. Certainly, the idea of personal choice emerged in Alex’s persuasive argument on this topic – so maybe it was. And yet, I felt more confident in coding this as a connection to background knowledge since there was no evidence that Alex was reading about vegans at the time or that the insight necessarily lead to a new understanding of the topic. A second round of analyses will return to these data so that the specific conditions under which this type of statement occurred can be described and traced to evidence of generative synthesis in essays. This analysis could inform developmental descriptions of generative synthesis for novice online readers.

On this point, students in this study did spend time constructing an understanding of individual texts (*M*) and, to a lesser degree, updating their understanding of their topic (*J*), but “schemas of the moment” that could be considered “generative” (DeSchryver, 2012) were a very rare exception. I think this is an important point. The ways that adolescent students use the

Internet, in school, while reading on curricular topics may represent an early moment in the learning trajectory toward more expert synthesis processes. Alternatively, the ways that these students have learned to construct meaning in school may actually have shaped their approach to online reading so that they prioritized the construction of meaning within texts (i.e., Identify Important Information) more than the construction of meaning across texts. Here, I draw on assumptions that reading and learning are inextricably rooted in context (Brown, Collins & Duguid, 1989; Lave & Wenger, 1991; Vygotsky, 1978). To nurture development of generative synthesis, as described by DeSchryver (2012), do students need different models of instruction, and explicit focus on integration earlier in their K-12 schooling? The CCSS articulate this expectation, but these are empirical questions that warrant further study.

### **Limitations**

Several limitations constrain the generalizability of these findings. Sampling is the most significant and obvious issue. A larger, more diverse sample would permit more robust conclusions. The limited results reported here reflect low power, which increased the threat of type II error. With more participants, additional effects, particularly with regard to within-group and between-groups application of  $[(PST)^2 + (iC^3)]$  strategies may have been detected. Future studies will address this limitation by extending the sampling frame, but also, as necessary, the length of time for data collection.

Given the constraints of the design, students in this study were only given three opportunities to practice with LINKS. One question that emerges is whether or to what extent LINKS, delivered over the span of a whole unit of study, or over the span of a semester, or even a whole year might impact the complex set of skills it was designed to scaffold. In some fashion, the limited impact of the intervention may have been a factor of too little time on task. Dwyer

(2010) found significant impact of intervention over two years of study but my study lasted only two months. It will be important to extend the period of release so that it is more gradual and offers more guided support through the early stages of skill development. A longer time frame would enable a more robust examination of LINKS' potential to impact multiple, multimodal Internet text integration processes and products.

It is fair to question whether or to what degree my own close involvement with data analyses may have influenced the outcomes of the study. Although I was deeply committed to rigorous checks and balances on my applications of codes and evaluation of integration indicators in students' essays, it cannot be entirely ruled out that I introduced bias. That said, in a study that has tried to break new ground in an emerging field of study, I needed to be intimately connected to the data in ways that a more objective distance would not have permitted.

It is important to note that a technical glitch that went undetected during the posttest meant that Dyad 4's online inquiry processes were not recorded. The recording stopped due to a technical issue after only 6 minutes of reading. I therefore used a mean value estimate of Dyad 4's posttest scores, calculated by using scores from their three practice sessions. This certainly introduced variability, but given my observation that this dyad consistently applied the same general approach to their work, I felt this choice was justified.

Variability was also introduced because of very real-world interruptions to the schedule of data collection. Student illnesses, scheduling conflicts due to tests, exams, and field trips delayed data collection to after the winter break. This was not part of the initial plan. The disruptions may have influenced the impact of the intervention for the participants. The upside, however, is that despite these very real sources of variability, the LINKS intervention was found to have a statistically significant impact on evidence of integration in students' persuasive essays.

Finally, the use of non-validated instruments also limits the claims that can be made in this study. This choice was taken because, so far as I am aware, validated instruments for the measure of integration in students' online inquiry and written work have not yet been published. One criticism would be that the design of the TII Index rubric was very closely aligned with the intervention so that the statistically significant results may, in fact, be more a result of this alignment than the actual impact of the LINKS intervention on students' ability to construct an integrated understanding from multiple, multimodal Internet texts. The rubric's constituent criteria may not, in fact, be the best cognitive trace indicators of multiple text integration and I concede that this is a fair criticism. That said, the risk taken here to develop methods of measuring these complex sets of processes offer a starting place for further inquiry. In this way, the rubrics, though certainly flawed, may be seen as a significant preliminary contribution to the literature.

## CHAPTER 6

### Conclusion

This study offers three contributions to the field of digital literacies intervention research. First, and notwithstanding the limitations outlined in the previous section, this study contributes preliminary evidence that a strategically designed and carefully choreographed intervention can nudge adolescents toward the construction of an integrated understanding of information gathered from multiple, multi-modal Internet texts during online inquiry, over and above simply practicing this skill without instructional support. As noted, even evidence of a “nudge” at this early moment in this field of research can be considered important. As teachers in the US adopt the Common Core State Standards (CCSS) for English Language Arts and Literacy in History/Social Studies, Science and the Technical Subjects (National Governor’s Association Center for Best Practices & Council of Chief School Officers, 2010) in the broader context of a digital age, it is essential for researchers and practitioners to carefully examine methods of instruction that support development of multiple, multimodal Internet text integration skills for all K-12 learners. These data provide hints of pedagogical promise.

Second, grounded in theoretical foundations of online reading, inquiry and instruction, this study provides the LINKS [Learning to Integrate InterNet Knowledge Strategically] intervention for teaching ninth grade students how to construct an integrated understanding of curricular topics through the finding and reading of multiple, multi-modal Internet texts. The intervention included seven integrated parts: (a) dyadic discussion of reading prompt, reading purpose and background knowledge; (b) quick, direct introduction and review of  $[(PST)^2 + (iC^3)]$  strategies and questions; (c) teacher modeling of strategy use for the purpose of constructing an integrated understanding of topics from multiple texts via a series of three screencasts that

gradually released responsibility to students over three intervention sessions, (4) 30 minutes of dyadic online inquiry; (d) guided teacher questioning that prompted application of  $[(PST)^2 + (iC^3)]$  strategies; (e) note taking that required students to change ink color to delineate information gathered from different information sources; (f) writing a persuasive argument for 20 minutes. It also includes a set of ten strategies,  $[(PST)^2 + (iC^3)]$  that could be considered the content, or “the what” of integration strategies instruction. This framework bundles strategies that have received support in the literature and aligns them with guiding questions that teachers and students can use. Although other strategies may warrant inclusion, as outlined,  $[(PST)^2 + (iC^3)]$  provides a preliminary framing for intervention research that, like this study has done, aims to support the development of multiple, multimodal Internet text integration skills. The coding scheme, aligned closely with  $[(PST)^2 + (iC^3)]$ , articulates a range of strategic activities within the top-level categorical codes that teachers and researchers may observe in similar contexts and for similar curricular inquiry tasks [See Appendix F]. The think aloud screencasts that I created for this study are all publicly available [See Appendix D].

Finally, this study provides a rubric for evaluating trace indicators of integration in students’ written arguments and a method for justifying the provenance of those trace indicators when articulated in students’ written arguments. Although open for refinement and critique, this instrument is my first attempt to fill a gap in the broader field of online reading comprehension research that could enable a more nuanced interpretation of integration for this particular type of discipline-specific online inquiry task—a type of task that is increasingly expected in schools and across disciplines (CCSS, 2010). Conversations around the criteria included in this rubric could lead to (a) the refinement of definitions of multiple, multimodal Internet text integration processes in online, school-based inquiry activities, and (b) enable the development of baseline

understandings of developmental trajectories for these skills across grade levels.

In sum, the LINKS intervention shows some promise of effectiveness and these findings raise many more questions than they answer. Having examined the question of impact (if any) in this study, future research should focus more specifically on how and why. A stronger understanding of the social, cognitive and contextual interactions at play could inform nuanced pedagogical practices and enable richer connections to theories of learning and adolescent development. In a networked world that offers unlimited access to more information and more perspectives than ever before, the ability to construct an integrated understanding from what is found and read on the Internet, is fundamentally important to every facet of the literate lives our students will lead. I hope this research will enable teachers and researchers to engage in productive conversations about online inquiry, multiple text integration skills development and to research that further elucidates the complexities of teaching adolescents to construct an integrated understanding from multiple, multimodal Internet texts.

## APPENDICES



## Appendix A: ACT Persuasive Writing Scoring Guidelines

These are the descriptions of scoring criteria that trained raters will follow to determine the score (1–6) for each essay.

### **Score = 6**

**Essays within this score range demonstrate effective skill in responding to the task.**

The essay shows a clear understanding of the task. The essay takes a position on the issue and may offer a critical context for discussion. The essay addresses complexity by examining different perspectives on the issue, or by evaluating the implications and/or complications of the issue, or by fully responding to counter-arguments to the writer's position. Development of ideas is ample, specific, and logical. Most ideas are fully elaborated. A clear focus on the specific issue in the prompt is maintained. The organization of the essay is clear: the organization may be somewhat predictable or it may grow from the writer's purpose. Ideas are logically sequenced. Most transitions reflect the writer's logic and are usually integrated into the essay. The introduction and conclusion are effective, clear, and well developed. The essay shows a good command of language. Sentences are varied and word choice is varied and precise. There are few, if any, errors to distract the reader.

### **Score = 5**

**Essays within this score range demonstrate competent skill in responding to the task.**

The essay shows a clear understanding of the task. The essay takes a position on the issue and may offer a broad context for discussion. The essay shows recognition of complexity by partially evaluating the implications and/or complications of the issue, or by responding to counter-arguments to the writer's position. Development of ideas is specific and logical. Most ideas are elaborated, with clear movement between general statements and specific reasons, examples, and details. Focus on the specific issue in the prompt is maintained. The organization of the essay is clear, although it may be predictable. Ideas are logically sequenced, although simple and obvious transitions may be used. The introduction and conclusion are clear and generally well developed. Language is competent. Sentences are somewhat varied and word choice is sometimes varied and precise. There may be a few errors, but they are rarely distracting.

### **Score = 4**

**Essays within this score range demonstrate adequate skill in responding to the task.**

The essay shows an understanding of the task. The essay takes a position on the issue and may offer some context for discussion. The essay may show some recognition of complexity by providing some response to counter-arguments to the writer's position. Development of ideas is adequate, with some movement between general statements and specific reasons, examples, and details. Focus on the specific issue in the prompt is maintained throughout most of the essay. The organization of the essay is apparent but predictable. Some evidence of logical sequencing of ideas is apparent, although most transitions are simple and obvious. The introduction and conclusion are clear and somewhat developed. Language is adequate, with some sentence variety.

and appropriate word choice. There may be some distracting errors, but they do not impede understanding.

**Score = 3**

**Essays within this score range demonstrate some developing skill in responding to the task.**

The essay shows some understanding of the task. The essay takes a position on the issue but does not offer a context for discussion. The essay may acknowledge a counter-argument to the writer's position, but its development is brief or unclear. Development of ideas is limited and may be repetitious, with little, if any, movement between general statements and specific reasons, examples, and details. Focus on the general topic is maintained, but focus on the specific issue in the prompt may not be maintained. The organization of the essay is simple. Ideas are logically grouped within parts of the essay, but there is little or no evidence of logical sequencing of ideas. Transitions, if used, are simple and obvious. An introduction and conclusion are clearly discernible but underdeveloped. Language shows a basic control. Sentences show a little variety and word choice is appropriate. Errors may be distracting and may occasionally impede understanding.

**Score = 2**

**Essays within this score range demonstrate inconsistent or weak skill in responding to the task.**

The essay shows a weak understanding of the task. The essay may not take a position on the issue, or the essay may take a position but fail to convey reasons to support that position, or the essay may take a position but fail to maintain a stance. There is little or no recognition of a counter-argument to the writer's position. The essay is thinly developed. If examples are given, they are general and may not be clearly relevant. The essay may include extensive repetition of the writer's ideas or of ideas in the prompt. Focus on the general topic is maintained, but focus on the specific issue in the prompt may not be maintained. There is some indication of an organizational structure, and some logical grouping of ideas within parts of the essay is apparent. Transitions, if used, are simple and obvious, and they may be inappropriate or misleading. An introduction and conclusion are discernible but minimal. Sentence structure and word choice are usually simple. Errors may be frequently distracting and may sometimes impede understanding.

**Score = 1**

**Essays within this score range show little or no skill in responding to the task.**

The essay shows little or no understanding of the task. If the essay takes a position, it fails to convey reasons to support that position. The essay is minimally developed. The essay may include excessive repetition of the writer's ideas or of ideas in the prompt. Focus on the general topic is usually maintained, but focus on the specific issue in the prompt may not be maintained. There is little or no evidence of an organizational structure or of the logical grouping of ideas. Transitions are rarely used. If present, an introduction and conclusion are minimal. Sentence structure and word choice are simple. Errors may be frequently distracting and may significantly impede understanding.

**No Score:** Blank, Off-Topic, Illegible, Not in English, or Void

## Appendix B: Trace Indicators of Integrative (TII) Index Rubric

<b>Criterion</b>	<b>Score</b> <b>0 = no</b> <b>1 = somewhat, or</b> <b>one example;</b> <b>meets minimum</b> <b>2 = yes definitely,</b> <b>or more than one</b> <b>example; exceeds</b> <b>minimum</b>	<b>Evidence/Justification</b>
Does the persuasive essay make an argument consistent with the expectations outlined in the topic prompt.		
Does the persuasive essay include information learned from more than one source?		
Does the persuasive essay include information learned from more than one medium?		
In the persuasive essay, is the central argument/position grounded in corroborating facts from two or more websites/texts?		
Does the persuasive argument include counterpoints to the central argument collected from one or more sources different from the sources used to construct the central argument?		
Does the persuasive essay integrate facts that were recorded as part of the author's bank of pre-existing knowledge?		
Does the essay provide evidence for construction of an integrated mental model of understanding: Is there evidence of integration of information across texts and/or within texts, and/or with background knowledge?		
Does the persuasive essay include linguistic markers indicative of integration (e.g., seriation, transitional phrases that connect ideas, connectives, parallel structures that show an integrated understanding)		

### Trace Indicators of Integration Rubric (cont'd)

Does the persuasive essay include explicit reference to source information [i.e. mention of author, a reason for why we should trust this information]?		
Does the persuasive essay include a thesis/synthesis statement that communicates an integrated understanding of the topic?		
	TOTAL:	

## Appendix C: Screencast Transcripts

At each stage of the intervention, I created a screencast for each topic, and for each condition. This was because topics were counterbalanced for participants and the intervention changed at each session according to the plan for gradual release. The transcripts provided here are for the third practice session for both treatment and control conditions on each topic.

### **Prompt 1**

#### **Wind vs. Coal**

##### **Treatment group**

Site 1

[http://www.ucsusa.org/clean\\_energy/coalvswind/](http://www.ucsusa.org/clean_energy/coalvswind/)

Site 2

<https://www.youtube.com/watch?v=FpucONE7WWk>

Welcome to your third and final screencast of the study. The purpose of these screencasts has been to help you develop skills and strategies to synthesize, or integrate, what you read from multiple online texts, and also with your pre-existing knowledge.

The PST2 + iC3 formula provides a list of strategies that you can use to build an integrated understanding of an issue as you locate information, evaluate it, synthesize it and update your understanding. Recall that last time, we focused exclusively on comparing information with pre-existing knowledge and on connecting information across information sources. With each bit of information you found, you asked yourself three questions -- 1) Is this the same as something I know or have read elsewhere? 2) Does it contradict or go against something that I knew or read elsewhere? and thirdly, is this new information that takes my understanding in a new direction or deepens it in some way.

This time, you will watch me and listen as I engage in my own online reading process in response to the prompt on wind vs. fossil fuels for energy production. This time, however, I won't be quite as explicit about naming the strategies I am using. Your task will be to watch, listen and to identify the strategies I'm using along the way. Also, as you identify the strategies, ask yourself if you would make the same choices as I have made.

Let's get started.

Okay, so I need to determine the relative environmental impact of wind vs. fossil fuels for power generation.

Here are my notes -- I've written down what I know already here. [show Google Doc with notes]

Search [go to Google, enter search terms & say aloud]: "fossil fuels vs. wind for energy production"

Hmmm...I'm skimming and scanning [on the SERP]. I'm going to pick this one. It's from a non-profit and it seems to confirm what I know already.

It's an informational website.

Who are these folks?

They're an advocacy group -- okay -- so this helps me to understand their perspective.

I see this information about wind.

I think I'll copy it and add it to my notes.

Okay -- so now that I have it in my notes, I'm thinking about how it compares with my background knowledge. It seems to confirm my pre-existing knowledge. Wind energy production doesn't pollute the air, water or land.

It also extends my understanding because I now know how it doesn't pollute -- which explains why it's considered clean energy.

I didn't know that fossil fuels were required to make wind turbines -- so that's completely new information, and it contrasts with my pre-existing knowledge. I thought wind was emissions free - but it may not be.

Also, I didn't know that birds are harmed.

Okay -- so what do I still need to understand? I'd like to see what other organizations say about wind vs. fossil fuels.

Go back to search results.

Is there anything here that I could use?

Hmmm...this video seems promising. It's short.

Who are the authors?

And, it's from an organization that seems to be advocating for the fossil-fuel industry which makes me wonder about the weight I should give their perspectives. There may be some truth...but the fossil fuel industry stands to lose a lot of money if wind power generation takes over...

I'm going to watch it.

Watch video.

Okay -- so what information seems important from that video?

1) Fossil fuels are used to build, transport and lubricate parts of wind turbines

This means that wind power is not emissions free, like I thought. It also seems consistent with what the Union of Concerned Scientists said. So, that's a connection across sources that seems to back up the same idea.

2) Subsidies on wind make energy more expensive for everyone. This seems like it could be a politically motivated statement. I think I will need to investigate this claim a little further.



I also wonder about the long-term environmental impact of emissions from the production of wind energy. If we build lots of turbines now, would the clean generation of energy in the future be worth the output costs now? Hmmm...lots more to research.

Okay -- so now, it's your turn.

Before you start reading, take a moment to identify the strategies that you saw me use as I read and thought aloud.

Also, would you have made the same choices? Did you make the same connections?

Share your thoughts with your partner and then, start your reading.

### **Control Group**

Welcome to your third and final screencast of the study. The purpose of these screencasts has been to provide you with some starter texts and information that you can use as you synthesize, or integrate, what you read from multiple online texts.

This screencast will help you get started with today's online reading and synthesis task.

As you know, you will need to determine whether windmills are a better way to make energy than power plants that use fossil fuels to make electricity. Then, you will need to construct a persuasive argument and use multiple Internet texts to do so because no one source could possibly tell you everything that you need to know on this issue.

Today, I'm going to search for a couple of sources that I could use to construct a persuasive argument to this question.

I'll pause so that you can read and listen silently. Then, like before, you'll get started with your own Internet reading and research.

Ready?

Enter search term [in Google]: "fossil fuels vs. wind for energy production"

Skim, scan down...and choose the Union of concerned scientists page.

Pause for same amount of time as in treatment video.

Return to list and pick the video.

Okay -- now that you have some information to get you started, it's your turn to read more.

Happy Reading.

## **Prompt 2**

### **Nuclear Fission**

#### **Treatment**

Welcome to your third and final screencast of the study. The purpose of these screencasts has been to help you develop skills and strategies to synthesize, or integrate, what you read from multiple online texts, and also with your pre-existing knowledge.

The PST2 + iC3 formula provides a list of strategies that you can use to build an integrated understanding of an issue as you locate information, evaluate it, synthesize it and update your understanding. Recall that last time, we focused exclusively on comparing information with pre-existing knowledge and on connecting information across information sources. With each bit of information you found, you asked yourself three questions -- 1) Is this the same as something I know or have read elsewhere? 2) Does it contradict or go against something that I knew or read elsewhere? and thirdly, is this new information that takes my understanding in a new direction or deepens it in some way. [put this at top of NOTES]

This time, you will watch me and listen as I engage in my own online reading process in response to the prompt on whether the peaceful use of nuclear fission for power generation are worth the risks. This time, however, I won't be quite as explicit about naming the strategies from the PST2 + iC3 formula that I am using. Your task will be to watch, listen and to identify the

strategies I'm using along the way. Also, as you identify the strategies, ask yourself if you would make the same choices as I have made.

Let's get started.

Okay, so I need to determine the pros and cons of nuclear fission.

Here are my notes -- I've written down what I know already here. [show notes in GDoc.]

Search [enter search terms in Google]: "risks vs. benefits of nuclear power"

Hmmm...I'm skimming and scanning. I'm going to pick this one. It's from a non-profit and it seems to confirm what I know already.

It's an informational website.

Who are these folks?

They're a group of scientists who say they have no vested interest in nuclear power beyond just providing facts.

I see this information in the summary.

I think I'll copy this to my notes.

Also this.

Okay -- so now that I have this information in my notes, I'm thinking about how it compares with my background knowledge.

It seems to confirm but also to go against my pre-existing knowledge -- I thought nuclear plants were emissions free -- but actually, they have LOW, not no, emissions. The mining of uranium and its conversion into nuclear fuel generates CO<sub>2</sub>...so this takes my understanding in a new direction.

As I update my understanding, I feel like my understanding of this one aspect of nuclear fission is a little deeper. But I still have more questions about the risks. I know already about nuclear disasters like Chernobyl and Three-mile Island. Are there other risks?

Go back to search results.

Is there anything here that I could use?

Hmmm...this video seems promising. It's short.

Who is this organization? Climate central -- well they're not advocating a particular position. But they do want to influence policy. I need to keep this in mind as I weight what I learn here.

Okay, I'm going to watch it.

Watch video.

So, what information seems important from that video?

1) First -- the nuclear physicist did mention the LOW CO<sub>2</sub> emissions -- that's a connection to the nuclearinfo source -- both sources gave the same information, and both sources seem pretty objective -- so I think I can put some weight in this as I construct my understanding.

2) The cons -- well, I knew about disasters like Chernobyl -- so "catastrophic accident" backs up what I knew. I also knew about nuclear waste being a problem -- but not the weapons connection. That seems pretty important.

I think I'll investigate how likely it is that an accident could occur -- but also how easy it is to make nuclear weapons from nuclear power generation technology. This seems like it would be important for my argument.

Okay -- so now, it's your turn.

What strategies did you identify as I read and thought aloud?

Would you have made the same choices? Did you make the same connections?

Share your thoughts with your partner and then, start your reading.

### **Control Group**

Welcome to your third and final screencast of the study. The purpose of these screencasts has been to provide you with some starter texts and information that you can use as you synthesize, or integrate, what you read from multiple online texts.

This screencast will help you get started with today's online reading and synthesis task.

As you know, you will need to determine whether the peaceful use of nuclear fission for power generation is worth the risks. Then, you will need to construct a persuasive argument and use multiple Internet texts to do so because no one source could possibly tell you everything that you need to know on this issue.

Today, I'm going to search for a couple of sources that I could use to construct a persuasive argument to this question.

I'll pause so that you can read and listen silently. Then, like before, you'll get started with your own Internet reading and research.

Ready?

Search term: "risks vs. benefits of nuclear power"

Skim, scan down...and choose the nuclear info.net page.

Pause for same amount of time as in treatment video.

Return to list and pick the video.

Okay -- now that you have some information to get you started, it's your turn to read more.

Happy Reading.

### **Prompt 3**

#### **Math vs. Chemistry Careers**

##### **Treatment**

Welcome to your third and final screencast of the study. The purpose of these screencasts has been to help you develop skills and strategies to synthesize, or integrate, what you read from multiple online texts, and also with your pre-existing knowledge.

The PST2 + iC3 formula provides a list of strategies that you can use to build an integrated understanding of an issue as you locate information, evaluate it, synthesize it and update your understanding. Recall that last time, we focused exclusively on comparing information with pre-existing knowledge and on connecting information across information sources. With each bit of information you found, you asked yourself three questions -- 1) Is this the same as something I know or have read elsewhere? 2) Does it contradict or go against something that I knew or read elsewhere? and thirdly, is this new information that takes my understanding in a new direction or deepens it in some way. [put this at top of NOTES]

This time, you will watch me and listen as I engage in my own online reading process in response to the prompt on whether advanced study of math or chemistry offers better career prospects. This time, however, I won't be quite as explicit about naming the strategies from the PST2 + iC3 formula that I am using. Your task will be to watch, listen and to identify the strategies I'm using along the way. Also, as you identify the strategies, ask yourself if you would make the same choices as I have made.

Let's get started.

Okay, so I need to determine whether advanced study of math or chemistry will lead to better job prospects.

Here are my notes -- I've written down what I know already here. [show notes]

Search: "math vs. chemistry careers"

Hmmm...I'm skimming and scanning. I'm going to pick this one. It's a dot org.

Pick weusmath.org/chemist

It's an informational website.

Who are these folks?

Oh - it's a website created by the math department at Brigham Young University. I feel like I can trust this information.

Okay -- so, here's some information I could use.

I think I'll copy this to my notes.

Also this.

Okay -- so now that I have this information in my notes, I'm thinking about how it compares with my background knowledge.

I knew that a lot of chemistry was required in certain industries, though not necessarily this many. I did know that chemists helped in the manufacture of many products like soap and plastic, so this confirms my pre-existing knowledge.

I didn't, however, know how much chemists can expect to make. This is an important consideration for future career prospects.

I also didn't know how much math was involved with chemistry.

Now I'm wondering if there are more job prospects with math specifically because it could be a foundational skill set for a lot of other jobs.

Let's see if I can find out more.

Go back to search results.

Is there anything here that I could use?

Hmmm...this video seems promising. It's short.

Who is this organization? wonderville.ca -- okay -- they seem to be a government organization that promotes science information -- that seems pretty reliable.

Okay, I'm going to watch it.

Watch video.

So, what information seems important from that video?

- 1) Chemical engineers use math! That backs up what I read at the other site.
- 2) Chemical engineers solve problems -- like how to turn plants into textiles. That's interesting. I didn't know that before.

Now, I realize that math and chemistry are connected, so I'm going to need to tell 9th graders about this in my argument. I also need to answer some more questions.

I don't know much about the number of jobs in math or chemistry fields - or how well math jobs pay relative to chemistry jobs. I need to find this out next.

Okay -- so now, it's your turn.

What strategies did you identify as I read and thought aloud?

Would you have made the same choices? Did you make the same connections?

Share your thoughts with your partner and then, start your reading.

### **Control Group**

Welcome to your third and final screencast of the study. The purpose of these screencasts has been to provide you with some starter texts and information that you can use as you synthesize, or integrate, what you read from multiple online texts.

This screencast will help you get started with today's online reading and synthesis task.



As you know, you will need to determine whether advanced study in math or chemistry offers better job prospects. Then, you will need to construct a persuasive argument and use multiple Internet texts to do so because no one source could possibly tell you everything that you need to know on this issue.

Today, I'm going to search for a couple of sources that I could use to construct a persuasive argument in response to this question.

I'll pause so that you can read and listen silently.

Then, like before, you'll get started with your own Internet reading and research.

Ready?

Search term: "math vs. chemistry careers"

Skim, scan down...and choose the [weusemath.org/chemist](http://weusemath.org/chemist) page

Pause for same amount of time as in treatment video.

Return to list and pick the video.

Okay -- now that you have some information to get you started, it's your turn to read more.

Happy Reading.

## Appendix D: Screencast URLs and Timings

Hyperlinks to all screencasts are curated at <http://mschirahagerman.com/downloads-and-links/>

This table also provides a listing of URLs. Please note that screencasts were recorded in HD using Camtasia (Techsmith, 2012). YouTube, however, adjusts the quality of the video in real time according to bandwidth demands. Quality of the video may, therefore, vary depending on when you view it. This did occur during the study. This meant that on occasion, I reverted to the recorded version of the video on my computer hard drive to ensure the highest quality viewing experience for all participants.

Table D1

### *Summary of Screencast URLs*

Topic	Session	Treatment Version	Control Version
Wind v. Coal	1	<a href="http://youtu.be/u_3b7BzUIzA">http://youtu.be/u_3b7BzUIzA</a>	<a href="http://youtu.be/5y5wKIkJW1w">http://youtu.be/5y5wKIkJW1w</a>
	2	<a href="http://youtu.be/-oY4DC0Pyko">http://youtu.be/-oY4DC0Pyko</a>	<a href="http://youtu.be/SCMj7ytaY9M">http://youtu.be/SCMj7ytaY9M</a>
	3	<a href="http://youtu.be/WRhk-UU-sVo">http://youtu.be/WRhk-UU-sVo</a>	<a href="http://youtu.be/1AijqEyocCw">http://youtu.be/1AijqEyocCw</a>
Nuclear Energy	1	<a href="http://youtu.be/Oz3fTMEb2gc">http://youtu.be/Oz3fTMEb2gc</a>	<a href="http://youtu.be/tXv11t5X1lg">http://youtu.be/tXv11t5X1lg</a>
	2	<a href="http://youtu.be/AxG1ARz32Bg">http://youtu.be/AxG1ARz32Bg</a>	<a href="http://youtu.be/QCaD76kS6AI">http://youtu.be/QCaD76kS6AI</a>
	3	<a href="http://youtu.be/_zKO2eraIGg">http://youtu.be/_zKO2eraIGg</a>	<a href="http://youtu.be/YXY1ILcfzgI">http://youtu.be/YXY1ILcfzgI</a>
Math v. Chemistry	1	<a href="http://youtu.be/uyo1jyGIB38">http://youtu.be/uyo1jyGIB38</a>	<a href="http://youtu.be/uyo1jyGIB38">http://youtu.be/uyo1jyGIB38</a>
	2	<a href="http://youtu.be/MN4Ez3_NpKE">http://youtu.be/MN4Ez3_NpKE</a>	<a href="http://youtu.be/DBSX3a_H85c">http://youtu.be/DBSX3a_H85c</a>
	3	<a href="http://youtu.be/sCmr2B2jPl4">http://youtu.be/sCmr2B2jPl4</a>	<a href="http://youtu.be/cRj6-xfu1tg">http://youtu.be/cRj6-xfu1tg</a>

Table D2

*Summary of Screencast Content and Timings*

Topic & Texts	Session	Content Time (Website + Video)	Treatment Screencast Length	Control Screencast Length
<b>Wind v. Coal</b>	1	3:10 + 1:25	14:55	9:29
Union of Concerned Scientists				
Website: <a href="http://www.ucsusa.org/">http://www.ucsusa.org/</a>	2	:46 + 1:25	8:44	3:06
Institute for Energy Research, Wind = Fossil Fuels Video: <a href="http://youtu.be/FpucONE7WWk">http://youtu.be/FpucONE7WWk</a>	3	:55 + 1:25	9:10	5:07
<b>Nuclear Energy</b>	1	3:11 + :46	15:18	10:52
Nuclearinfo.net website <a href="http://nuclearinfo.net/">http://nuclearinfo.net/</a>	2	1:10 +:46	7:41	2:34
What are the benefits and risks of nuclear power? Climate Central <a href="http://youtu.be/h9xFcoou1f8">http://youtu.be/h9xFcoou1f8</a>	3	1:20 +:46	7:48	5:11
<b>Math v. Chemistry</b>	1	3:49 +2:23	17:34	9:49
Brigham Young University Math Department Website <a href="http://weusemath.org/">http://weusemath.org/</a>				
Wonderville.ca Cool Careers: Chemical Engineers <a href="http://youtu.be/_UXwbxM8YfI">http://youtu.be/_UXwbxM8YfI</a>	2	1:10 + 2:23	8:51	4:09
	3	1:20 +2:23	9:22	6:22

## Appendix E: Topic Prompts

### **PRE-POST Prompt 1**

**Curriculum Expectation:** Propose possible effects (on the genes) of exposing an organism to radiation and/or toxic chemicals. (B4.2E)

All mothers and fathers want to ensure their babies live in a safe and healthy environment but, understanding the potential impact of common household and baby products on a baby's development takes specialized knowledge. What information do parents need in order to make good choices for their families?

Using multiple, trustworthy Internet texts of any type (e.g., print, photos, video, graphics, charts, figures, tables etc.), read about possible effects of chemicals commonly found in soaps on child development. Then, using what you have learned, write a persuasive argument for parents of young children that is for or against the use of synthetic soaps.

Remember also, to talk and think aloud as you read. You can jot down notes using any method you choose. You are allowed to use video, audio and pictures as information sources in addition to text. You will have time at the end of your reading session to talk about what you've learned. Then, you will write your argument.

### **PRE-POST Prompt 2**

**Curriculum Expectation:** Propose possible effects (on the genes) of exposing an organism to radiation and/or toxic chemicals. (B4.2E)

Doctors prescribe radiation treatment for certain kinds of cancers. Radiation treatments kill cancer cells by interfering with their DNA so that they can no longer reproduce. However, radiation therapy can also impact healthy cells and healthy DNA in our bodies. If someone has cancer, should they accept the risks?

Using multiple, trustworthy Internet texts of any type (e.g., print, photos, video, graphics, charts, figures, tables etc.), read about the possible effects of radiation treatment on cancer cell DNA versus healthy DNA in healthy cells. Then, using what you have learned, write a persuasive argument for someone with cancer that is either for or against radiation treatment.

Remember also, to talk and think aloud as you read. You can jot down notes using any method you choose. You are allowed to use video, audio, pictures, graphs, charts, figures etc. as information sources in addition to text. You will have time at the end of your reading session to talk about what you've learned. Then, you will write your argument.

### **Practice Session Prompt 1**

**Curriculum Expectation:** Describe renewable and nonrenewable sources of energy for human consumption (electricity, fuels), compare their effects on the environment, and include overall costs and benefits. (E2.4A) □

Your family just took a road trip to Toronto from Michigan. As you drove along the highway, you saw fields and fields full of windmills that are generating “clean energy”. Clean energy is a big topic in Michigan these days as the state looks for new industries. Are windmills a better way to make energy than power plants that use fossil fuels to make electricity?

Using multiple, trustworthy Internet texts of any type (e.g., print, photos, video, graphics, charts, figures, tables etc.), read about the environmental impact of wind energy vs. coal-fired electric plants. Then, using what you have learned, write a persuasive argument on the environmental impact of wind vs. coal for energy production that would convince Michigan lawmakers of your position.

### **Practice Session Prompt 2:** □

**Curriculum Expectation:** Describe peaceful technological applications of nuclear fission and radioactive decay. (P 4.12A)

Anti-nuclear advocates say there are no safe uses of nuclear energy. However, many countries around the world use nuclear fission peacefully to meet their energy needs. Are the peaceful uses of nuclear fission important enough to outweigh the risks?

Using multiple, trustworthy Internet texts of any type (e.g., print, photos, video, graphics, charts, figures, tables etc.) read about the risks of nuclear fission and the peaceful uses of this technology. Then, using what you have learned, write a persuasive argument for leaders of a country considering nuclear power that would convince them of whether to use nuclear fission or not.

### **Practice Session Prompt 3**

**Curriculum Expectation:** Evaluate the future career and occupational prospects of science fields. (C1.2E.)

There are many factors to consider when choosing the best career path, but one important factor is the range of careers you could do with advanced study in given subjects like science or math. Your Science teacher has asked you to research the types of jobs you could do with training in chemistry vs. the types of jobs you could do with training in math. □

Using multiple, trustworthy Internet texts of any type (e.g., print, photos, video, graphics, charts, figures, tables etc), read about whether advanced study of chemistry or math offers better future job opportunities. Then, using what you have learned, write a persuasive argument that recommends advanced training in either math or chemistry to ninth-grade students around the country.

## Appendix F: List of Codes

*Note.* Top-level categories are listed as headings in bold. Subcategorical codes are listed as bullet points. Letters were assigned arbitrarily to top-level categories using Morae (Techsmith, 2012) which constrains coding choice to letters of the alphabet.

### **A: Purpose**

- Using purpose to plan online inquiry process, especially at the start of the reading session

### **B: Pre-existing Knowledge**

- Pre-reading discussion of what students already know on the topic

### **C: Search**

- Planning search terms through discussion and negotiation
- Typing/entering search term (new search terms)
- Revising or refining search terms
- Using auto-complete to pick a search term after entering key word stem
- Going to Google (entering [www.google.com](http://www.google.com) or by some other method)
- Planning an overall search strategy (e.g., let's search on this topic and then on this topic so we can compare)
- Tangential search to fix-up gap in understanding
- Expressing emotion -- frustration, dismay, general anxiety because student doesn't know what search terms to use, or because they needed help from Google
- Using search function within a website

### **D: Source Selection**

- Skimming and scrolling down and up a list of search results - with eyes and mouse
- Hovering with mouse; scrutinizing for potential utility
- Reading search result title, URL or part of snippet text aloud as a way to (a) identify a link as having been viewed and considered, (b) communicate this information to partner, (c) consider the link's utility.
- Previewing link using >> arrows in browser
- Making and articulating predictive inferences using clues from info in search result entry (i.e., snippet text, URL, Title, or the >>preview)
- Clicking on first search result without evidence of strategic inferencing beyond perceived content-based relevance
- Using in-site menus and/or suggested hyperlinks
- Entering a URL directly into address bar
- Returning to list of search results after reading a text; reviewing, reconsidering potential utility of these results
- Using previous knowledge to identify the potential utility of a site

**E: Type**

- Using knowledge of source structure or source genre to make predictive inference about the type of information the source will contain, the relevance of information and/or its reliability

**F: Trustworthiness BEFORE selecting and reading**

- Identifying or questioning authorship before selecting source
- Identifying or questioning site credentials [of any kind] before selecting source

**Y: Trustworthiness DURING reading**

- Questioning and/or critically evaluating trustworthiness of text or information source based on clues and information extracted from text
- Identifying and/or evaluating the relative weight that information from a given source should have based on its' assessed trustworthiness

**G: Identify important information**

- Reading aloud
- Reading silently
- Skimming and scrolling down the page to identify information to read
- Skimming the text while reading -- skipping to information that seems most relevant, important and/or useful
- Previewing/Sampling a text or a text section to determine its utility before reading closely
- Monitoring or Predicting the utility of a text once engaged in the reading process
- Taking turns reading from a text
- Explicitly Identifying facts as interesting, relevant, important, just what they need or not what they need
- Using "find" feature within a page or a .pdf online to locate a keyword and therefore information that is most relevant to purpose/reading question

**H: Compare to background knowledge**

- Stating explicitly how information in text compares with pre-existing knowledge (confirms, contradicts, extends)
- Articulating an affective response to a text with words and/or paralinguistic expression (e.g., huh? Wow) in a way that shows the student is surprised, bothered, intrigued, impressed because this is new information or information that confirms what he/she already knew

**I: Connect with other texts**

- Identifying and stating explicitly how ideas in current text connect to ideas in another text read during study (inter-textual connection)
- Identifying and stating explicitly how ideas in current text connect to ideas read earlier within text (intra-textual connection)
- Identifying and stating explicitly how ideas in current text connect to ideas in a text read outside of the study (extra-textual connection)
- Planning reading path in order to facilitate inter-textual connections/comparisons

**J: Continually Update**

- Re-reading the prompt, checking to be sure the prompt is understood or re-focusing on purpose in a way that informs or enables a next step
- Discussion that supports integration of ideas, connects to the purpose of their reading activities, and/or furthers their reading process
- Identifying what is known on the issue and/or what is not yet known as a way to clarify logical next steps
- Using a focus on purpose to prompt planning for next steps
- Reconsidering interpretations of an idea after reading

**K: Student Question**

- Questions that are of a procedural or technical nature or that ask for repetition/clarification of what was said; these are not questions that focus on the construction of meaning.

**M: Constructing Understanding within a single text**

- Summarizing, paraphrasing and/or restating to construct understanding of meaning
- Word-level and/or sentence-level monitoring/fix-up
- Asking question to clarify understanding of text

**N: Student Takes Notes**

- Student(s) record, in writing, what they identify as important from the text they are currently reading
- Monitoring/metacognitive commentary/evaluation of note-taking process

**Q: Researcher Question - Treatment Group**

- Questions asked of treatment students during practice/treatment sessions; designed to support learning or focus on the reading process in some strategic way

**S: Researcher Scaffolding - Treatment Group**

- Any comment, discussion, ideas shared with treatment students and that is designed to support students' multiple text integration processes in some way

**P: Researcher Check-in Control**

- Any question or comment that Michelle asks/provides to the control condition students during their "practice" sessions; designed to elicit information about what is being read and/or control for "teacher presence" in students' reading experiences

**L: Reading on Topic that is Tangential to the Reading Purpose**

- Any reading activity that takes the student off in a direction that is only very loosely connected to the prescribed reading purpose, as outlined in the prompt

**R: Bounce**

- Clicking to a site and then immediately clicking out, usually without indication of the reason



**O: Observation Worth Exploring Further**

- This is a catch-all code that permits me to go back to examine situations like (a) technical failures and how kids resolved them or (b) conversations/interactions that weren't directly related to the RQs but may be of relevance to another study

## Appendix G: Representative Excerpts for Each Coding Category

*Note:* Representative excerpts have been extracted from the transcripts of treatment and control dyads. ST = Search Term; (T) = Treatment Dyad; (C) = Control Dyad

AN, BK, BL, BS, BZ, CS, DT, FD, HU, LG, MO, PO, SA, SR, TN, WS are participants in the study.

### Examples of A: Purpose

1 (T)

[@ 2:00] ALEX: Alright so, synthetic soaps and babies. [restating the main ideas]

2 (C)

AN:[summarizing what their reading purpose is] Okay, so cancer cells dna vs. healthy dna

### Example of B: Pre-existing Knowledge

1 (T)

BL: Do you have any more background knowledge? I don't have any so, I'm just gonna write that.

DT: I'm not gonna go too verbal on this. [small talk]

BL: Okay, so I have no prior knowledge. How are we going to approach this. [a suggestion to plan search]

DT: We have prior knowledge. We know that radiation can kill healthy cells. I mean it kind of said that. BL: It's in the prompt. Fine, you can write that down. So, no prior knowledge.

2 (C)

AN: Well, children don't really use a specific type of soap. BS: Well, there's like baby soap.

Johnson's and Johnson's. AN: Ya. That's sort of more like, I don't know. Seven year-olds. BS: [Shakes her head.] Baby soap. AN: Baby powder.

### Examples of C: Search

1 (C)

[@ Google] AN: [types: cons] I mean, we found pros [meaning, during the silent reading time with the video] Maybe we should go for cons now? [ST: cons of nuclear fission]

2 (C)

BS: [types google.com] So now, we're looking up Chernobyl. [into search bar: chernobyl]

3 (T)

JOHN: [starts to type in search phrase] synthe-

ALEX: It's already there. [meaning that autocomplete has already suggested an appropriate search term]

JOHN: [chooses autocomplete suggestion for first search: synthetics materials in children soap]

ALEX: [...] You have to click search.

JOHN: That might be a good idea.

4 (T)

21:08 ALEX: [types in new ST: is synthetic soap cheaper]

21:38 ALEX: [Refines his ST: is synthetic soap cheaper than natural soap]

5 (C)

LG: I guess we could try going to a different site. [clicks on back button from [www.lesstoxicguide.ca](http://www.lesstoxicguide.ca) website to go to Google search results page]

LG: [at search results page] Um. Let's type in something else. Um. [dictating search phrase] effects of chemicals in baby soaps [MO types it in to search bar - Affects of chemicals in baby soaps] [LG sees notice at top of results page: Showing results for *Effects* of chemicals in baby soaps -- she clicks on this corrected version]

## Examples of D: Source Selection

1 (T)

[ST: harmful effects of natural soaps]

JOHN: [hovers over first result] Oh. Here's a good one. [Title: How do I Choose the Best Baby Soap? URL: [www.wisegeek.com/how-do-i-choose-the-best-baby-soap.htm](http://www.wisegeek.com/how-do-i-choose-the-best-baby-soap.htm)] ALEX: Hm, but. I don't [in a skeptical tone]

JOHN: But that's what we're ask - trying to find out, is the best soap for their child.

ALEX: [looks at prompt] I guess so.

JOHN: [clicks] We don't always have to write something down. We can just look at it.

2 (C)

[Considering menu options in a page that they're already read for additional texts of use.] [BZ clicks on Nuclear Materials from top-level menu and then selects Special Nuclear Material from sub-level menu.]

3 (C)

WS: Top 10 lists!

CS: [reads the title] Highest paying jobs. [the first result] [she scrolls to the second] High paying chemistry jobs -- we could do this one.

WS: Wait, no, that's with only bachelor's degree [reading the third title] There's some high paying options in the field of chemistry WS & CS [in unison] -- and it's a dot org. [also coded as F: Trustworthy]

CS: [clicks] [Title: What are some high-paying career options in the field of chemistry? URL: [degreedirectory.org/.../What\\_are\\_some\\_high-paying\\_career\\_...](http://degreedirectory.org/.../What_are_some_high-paying_career_...) Snippet: Would you like to mix chemicals and create products? If so, the opportunities in the field of chemistry can be very competitive. The High-paying...]

### **Examples of E: Type**

1 (C)

TN: [She clicked on Baby Synthetic Soap but TN recognized immediately that the site was a e-commerce] Oh no, that's just to buy it. [Then they clicked away.]

2 (T)

BL: It's a blog. [he clicks away]

### **Examples of F: Trustworthiness Before Selecting and Reading Closely**

1 (C)

SR: [before reading closely] Um, that may not be the most reliable source. [commenting on yahooanswers.com question about the cost of enriched uranium]

BZ: Well no. But generally, things like that are a good place to start.

[BZ clicks away from the site]

2 (T)

DT: Are you sure we can trust this site? Oh, actually, never mind [she reads the title]

BL: [laughs] National Cancer Institute

DT: I read that and said never mind

BL: Oh, you can't trust it - DOT GOV?

DT: You can. Don't be sassy.

3 (T)

[At first glance of site, after M suggests it looks like a good site.]

PO: Yes, this is the United States Department of Labor [using a very confident tone] Sounds promising.

### **Examples of Y: Trustworthiness During Close Reading**

1 (T)

BL: So, do we still trust this site?

DT: It's the same website, BL, it's just different articles.  
BL: I know, but. Hold on. Are these user contributed?  
DT: hmhm.  
BL: cause it might be a little biased.  
DT: How could it be biased?  
BL: I mean like.  
DT: How can it be biased-  
BL: We can keep this, we can look at this, but we should use it with caution.  
DT: How can you be biased about cancer?  
BL: Not necessarily against cancer, more biased against radiation treatment. Could be lobbyists.  
DT: They're gonna be like - let me finish -- people are going to be all against cancer --[inaudible, but generally, she makes comments about how it would be stupid to lobby against radiation treatment] they're not going to be that stupid, come on.  
BL: [laughs a little]

2 (C)

BS: This seems legit [then she grimaces in a way that suggests she was being sarcastic. She questions the reliability/value of this source while AN explores hyperlinks further]

3 (T)

CS: [having just recognized the inconsistencies in salary reporting between myplan.com and weusemath.org] I trust the other website more though. WS: Me too. It was a dot org and this isn't.

### **Examples of G: Identify important information**

1 (C)

BS: Wait [she's reading and scrolling to find info of relevance]  
[Title of article: People died at Three Mile Island - The Free Press - Independent URL: [freepress.org/columns/display/7/2009/1733](http://freepress.org/columns/display/7/2009/1733)]  
AN: Um Public lied to. The public was lied to and there was an explosion.

2 (C)

[@ [www.lesstoxic.ca/index.asp?fetch=babycare](http://www.lesstoxic.ca/index.asp?fetch=babycare)]  
MO: [reading silently] [they click on a subpage from the homepage that focuses on soap, specifically [www.lesstoxic.ca/index.asp?fetch=babycare#soap](http://www.lesstoxic.ca/index.asp?fetch=babycare#soap)]  
LG: Let's see. Ummm [reading silently]  
MO: We should probably write that down.  
LG: Okay. [ @3:46 They position their mouse over the introductory paragraph to this page - LG copies down the first list of toxic chemicals]

3 (T)

[They click on heading, What are the side effects, and start reading about the side effects]  
[strategic episode ends at 16:11]

BK: Okay. [clicks on the title What are the side effects]

PO: Side effects. Fatigue. [laughs]

BK: Red, itching, peeling skin [grimaces]

PO: Okay, side effects [taking notes] are red, itching, I say scratching when I mean itching, I say it scratches, peeling skin, in your treatment area, this usually happens after about four weeks.  
[noting] four weeks post-treatment.

BK: Treatments.

PO: Then, there's more. There's fatigue as a side effect.

BK: We have to write the paragraph.

PO: You may feel tired. Well, fatigue is obvious. [writing notes] you may feel more tired than usual

BK: I'm pretty sure everyone in America feels fatigue.

PO: What, you just said to write it down.

BK: I just wrote fatigue -- I didn't write you may feel more fatigue than usual.

PO: Well, you just told me to write it down. You're so annoying. Wait. Loss of appetite, ap-pa-pa-pa. Appetite. Hair loss.

BK: Okay.

PO: Bam. How do I take care of my skin? What to do. Leave the marks on your skin until treatment is finished. Go down. Will I be able to have sex? That's up to you. Okay, that's pretty obvious. What other things can I do to help myself? Eat well. [taking notes] To help. Eat healthy. Drink at least 8 cups of water...8 cups of H2O a day. Drink soap and milk? oh, soup. I was really confused. Get some exercise and plenty of rest. [noting] exercise and plenty of rest.

### **Examples of H: Compare to Background Knowledge**

1 (C)

BZ: [reading silently about Chernobyl]

[The text states: There have been serious accidents with a small number of nuclear power stations. The accident at Chernobyl (Ukraine) in 1986 led to 30 people being killed and over 100,000 evacuated. In the preceding years, another 200,000 people were resettled away from the radioactive area. Radiation was even detected over a thousand miles away in the UK as a result of the chernobyl accident. It has been suggested that over time 2500 people died as a result of the accident.]

BZ: Chernobyl, ya. You still can't walk on the grass there.

[The text includes NOTHING about not being able to walk on the grass. BZ made a connection to background knowledge.]

2 (T)

PO: I know we have natural soaps at the house, but they're like \$15 a bottle.

3 (T)

WS: I didn't know that Chernobyl was real. I thought it was just a movie thing.

### **Examples of I: Connect with Other texts**

1 (C)

[after reading about the nuclear disaster at Three Mile Island in the USA]

BS: Chernobyl was probably more popular AN: Because people died. Popular? [they laugh in that way that suggests they don't consider it the best choice of words...]

2 (C)

BZ: [nodding his head as though to acknowledge that he has read this before -- i.e., in the snippet text and at nrc.gov previously] 235, plutonium 239.

3 (T)

ALEX: There's more about the fragrances being bad to be exposed to. So, we know that's true.

JOHN: Where are you reading that? Oh. Right.

### **Examples of J: Continually Update**

1 (C)

BS: [after taking notes] I kind of want to see how many people are employed AN: At this place?

BS: No, at an average nuclear power plant. Like, to see if it would create jobs.

2 (T)

JOHN: [seemingly inspired by the last passage he read] Well, maybe that's a question we could ask. What are the effects of F, D and C colorants in soap dye.

3 (C)

BZ: [talking aloud] Okay, ya, so you can use Uranium 235. [This is VERY interesting. BZ has been wondering about this question since he was at the nrc.gov website and was blocked when trying to access info about nuclear materials. He sees Uranium 235 as an autocomplete option, which seems to confirm for him that you can use Uranium 235 in the production of nuclear weapons.]

4 (T)

PO: [reading from prompt to remind herself what the purpose is] Radiation treatment for certain kinds of cancers.

BK: [clicks the back button in browser]

PO: Go to, go to pictures.

BK: [clicks on images]

PO: Never mind. That's Gray's Anatomy. [laughter] Go to...  
BK: Okay, we're just going to go to cancer radiation treatment.

### **Examples of M: Constructing Understanding within a single text**

1 (T)

ALEX: So, this is a pro-soap site? [he's asking this question to construct an overall understanding of the author's stance, what the objectives of the site might be] Natural soap?

JOHN: Hmhm. [nods in agreement] Shea butter. That's natural.

ALEX: So, they're talking about the dyes of the soap?

JOHN: I think so. [...] Or things that may contribute to the soap. Content.

2 (T)

[reading from the [www.jobbankusa.com](http://www.jobbankusa.com) site] BK: So most of them make between \$56 000 a year and \$91 000 a year.

PO: 56, wait, where is this? [looking more closely at the text]

3 (C)

AN: can I go down a little? [navigating down the page]

AN: Okay, so most of the side effects go away in two months. [summarizing content she has read]

### **Examples of N: Student Takes Notes**

1 (C)

[BZ and SR both write down facts from the [www.technologystudent.com](http://www.technologystudent.com) site.]

While taking notes, BZ says aloud: Radioactive waste [...] Power, of course [...]

BZ: Uranium in long rods. Kept cool in water. Accidents such as Chernobyl. Waste remains radioactive for thousands of years. Weapons. Power is equivalent to fossil fuel. No polluting gases. No longer {want?} to use fossil fuels. Small.

SR: Uranium is not an extremely rare materials. Formed into rods, submerged into water.

Chernobyl accident - 2500 people died over time. Amount of energy equal to that of fossil fuel plants. Don't produce harmful gasses. Is an economic alternative to fossil fuels.

2 (T)

[WS & CS both take notes]

WS: So, it's well developed...technology we...cheap and reliable.

[WS writes: well developed, cheap and reliable]

[CS writes: well developed - used for a long time, cheap and reliable]



## **Examples of K: Student Questions**

1 (T)

ALEX: Where do you see that? JOHN: It's all in here. [question focused simply on where to find the information that HU identified as important]

2 (C)

BS: How do you spell pharmacist?

AN: How do you think? It's with an F.

BS: No it's not.

## **Examples of Q: Researcher Question [During Treatment Intervention]**

1 (T)

M: So, you're thinking about eHow? You weren't sure whether to pick it or not?

CS: We're just going to look for the jobs and then

WS: A biochemist, we can look up different -- so we don't have to trust what they say necessarily

M: But you can corroborate, right? You can compare it?

WS: Hmhm.

2 (T)

Michelle: so, it looks like you've built some comparisons with what you know? Do you feel like it's confirming? Or do you feel like you've gone in some new directions?

3 (T)

Michelle: So, you moved away from that one because?

CS: It didn't really focus on what we were looking for.

WS: It was talking more about like, This is good. You should do this, instead of actually saying why.

M: Ah. And so, in terms of your reading purpose, did you feel like it wasn't going to get you where you needed to go?

CS: [nods] Ya.

4 (T)

M: So, what kind of a site is this? One of the strategies you can think about is the kind of site it is, right? Which will help you to predict the kind of information you might find there.

5 (T)

M: So, what do you feel like you know already? So, you've got two more minutes. How are you going to use those two minutes? What do you still need to know?

6 (T)

M: How are you doing?

BL: We have nothing.

M: Okay. Let's be strategic. What information. How are you crafting your search terms?

BL: That's not. Maybe I should say

DT: Careers

M: Math vs. chemistry careers [emphasis on careers since they haven't searched on that at present] That could refine it a little bit better, right?

7 (T)

M: Okay, so, who, who, who. Okay, so let's look at our formula here.

BK & PO: Ya.

M: So, you picked this source. It looked promising because it wasn't Wikipedia and it said pros and cons. So, the next thing is type and trustworthy. So, what kind of a site is this and how trustworthy is it?

8 (T)

M: So when you check in with yourselves and update your understanding, can you identify an idea that you feel you would need in order to construct your argument? Do you feel that something's missing, or? [the girls don't answer, they do look at their notes, however and whisper below their breaths as though they are trying to figure out what they know, don't know and still need to know, or some combination of those ideas, anyway]

### **Examples of S: Researcher Scaffolding [Treatment Condition]**

1

M: You thought it was just what, WS?

WS: just that movies that they had come out

M: Oh, ya. No. It happened when I was a kid. It was horrible.

M: And you brought up though, in Japan, after the tsunami, you brought that point up yourself. I overheard you say it. They're similar. Well, different cause, but similar circumstances for the poor people who were affected by it.

2

M: That's something that's great. I like how you just skimmed over that and were like no, that's not going to give me what I need so you moved on. That's great. [scaffolding a focus on inferencing/prediction based on the title, URL and snippet text]

3

M: So if you go back to you, the purpose, you're going to try to provide information about the range of jobs you could do with advanced study of these subjects and you're going to try to give this information to ninth graders, like you, around the country. So, maybe one thing you could do here is get a general sense of the chemistry jobs that are available so that you could use those as a comparison or in some way give a sense of the range of careers and the prospects, like how much money you could make or how much opportunity there seems to be for these jobs. So that could be something to help focus your thinking on this website or other places. So, just in terms of our framework. Sometimes, I know when I'm reading I need to check in with myself because I can go down a path on the Internet with my searching and then I have to ask myself what is it that I need to accomplish? Sometimes that's a helpful strategy to help you decide what you want to do next.

### **Examples of P: Researcher Check-in [Control Group]**

1 (C)

M: How's it going?

TN & SA: Good.

2 (C)

M: Sixteen minutes left to read.

3 (C)

M: [stands there and quietly asks them to focus on the task]

AN: Sorry [Michelle nods]

4 (C)

M: What are you looking at now?

AN: I'm looking for the top 10 chemistry - or would it be better to look up science?

M: That's up to you.

AN: I'll do chemistry and science.

5 (C)

M: What are you thinking about now, ladies?

LG: I didn't realize that wind power plants affect the animal populations.

MO: Ya. That's what I was thinking.

LG: Also, and I didn't think about that before and how that affects them.

6 (C)

M: Are you having internet issues? Just being slow?

LG: Ya.

M: Okay.

7 (C)

M: [in response to Bz's comment that YahooAnswers is a good place to start] What one are you on?

BZ: We're looking at Answers dot Yahoo. It's usually a good place to start.

8 (C)

M: Remember, guys, to talk aloud as you're thinking about what you're reading today.

## Appendix H: Representative Essay Evaluations and Markups

*Note:* Annotations that identify the source of information for highlighted sections are provided in Section 2.

### Higher Scoring Essay, Control Participant, Practice Session 1

In light of ongoing conflict regarding nuclear power plants, many opinions are available and each a little bit different from the rest. For the most part, there are two main opinions regarding the use of nuclear power plants: they *should* be utilised and they *should not* be utilised. After conducting internet research, I would advise you to take full advantage of nuclear power.

Although there are positive and negative effects of nuclear power- as there are pros and cons of any idea or action to be made-in this case, the positives outweigh the negative. Among these positives are that we already know that nuclear power works- we have been using it since the 70's. Nuclear power generation emits the least amount of carbon dioxide and greenhouse gases of all methods of energy generation. It is easy to expand generation, and in addition to this, expansion will create about 1,400-1,800 jobs to build each plant (with 400-700 permanent jobs retained). Also, nuclear power plants were designed to have long stretches of energy generation to minimize "down time" for refueling.

A few negative ideas are joined at the hip to nuclear power: the threat of terrorism, and the haunting memories of accidents like Chernobyl and Three Mile Island. Due to the radioactive nature of the materials involved in nuclear power plants, there is worry that terrorists will steal things like plutonium 239 or uranium and use them in weapons. This is certainly a valid concern, but the plants would have security measures. As for hesitation towards nuclear power because of previous accidents, all lives lost are tragedy, and it

should not be taken lightly. That said, what do you have to do when you fall off your bike while learning to ride? You have to pick yourself up, dust yourself off, and try again.

Nuclear power is the answer to a lot of our current energy needs. We need a clean source of reliable energy, something that we know will work for a long time and that won't run out. Nuclear power fits all of these criteria as well as holding massive potential for job growth.

Table H1

*TH Rubric Evaluation of Higher Scoring Essay*

Criterion	Score 0 = no 1 = somewhat, or one example 2 = yes definitely, or more than one example	Evidence/Justification
Does the persuasive essay make an argument consistent with the expectations outlined in the topic prompt.	2	Yes
Does the persuasive essay include information learned from more than one source?	2	Yes
Does the persuasive essay include information learned from more than one medium?	0	Text and video from the screencast, but everyone used this, so it's not during their reading time.
In the persuasive essay, is the central argument/position grounded in corroborating facts from two or more websites/texts?	2	Yes - from the video (nuclearinfo.org) and fi.edu/guide/wester/benefits.html & jobs numbers from nei.org

Table H1 (cont'd)

Does the persuasive argument include counterpoints to the central argument collected from one or more sources different from the sources used to construct the central argument?	2	Yes -- she cites cons from wikipedia, nuclearfissioni.wikispaces.com, world-nuclear.org
Does the persuasive essay integrate facts that were recorded as part of the author's bank of pre-existing knowledge?	1	Not from bgk recorded before the video -- but yes, from the video
Does the essay provide evidence for construction of an integrated mental model of understanding: Is there evidence of integration of information across texts and/or within texts, and/or with background knowledge?	2	Yes.
Does the persuasive essay include linguistic markers indicative of integration (e.g., seriation, transitional phrases that connect ideas, connectives, parallel structures that show an integrated understanding)	2	In addition, also (use of connective to bring together ideas from more than one source)  It is easy to expand generation, and in addition to this, expansion will create about 1,400-1,800 jobs to build each plant (with 400-700 permanent jobs retained). Also, nuclear power plants were designed to have long stretches of energy generation to minimize "down time" for refueling.

Table H1 (cont'd)

Does the persuasive essay include explicit reference to source information [i.e. mention of author, a reason for why we should trust this information]?	0	No. She does mention research, but not specific sources.
Does the persuasive essay include a thesis/synthesis statement that communicates an integrated understanding of the topic?	2	Yes. Final paragraph.
	TOTAL:	16

### Lower Scoring Essay, Treatment Participant, Practice Session 1

Nuclear fission is a type of nuclear power that provides energy for an alternate energy use.

Nuclear power comes with many risks and disadvantages, although it does have positive effects as well.

Everything has both pros and cons, and nuclear power is no exception to this rule. The pros of nuclear power include the low CO2 emissions, the amount of energy it provides (200 times that of burning coal) and its low operation expense. Though its pros sound like good things the cons of nuclear energy far outweigh the good. Nuclear power plants take up a lot of space and are very expensive to build, although once built they are inexpensive to maintain. This type of power comes with many dangers, some of which are- the radiation waste is very hard to get rid of safely and is extremely harmful to both humans and the environment; the workers have a high-risk of a meltdown within the power plant that could lead to a nuclear explosion; the center could be the target of a terrorist attack, because of



the damage a nuclear explosion can cause, and fission bombs (also known as atomic bombs) are a possible result from this process.

The radiation waste from nuclear fission plants is extremely dangerous, and hard to store safely, this waste lasts for 200 to 500 years . If an atomic bomb was created as a result of the power plant the effects could be devastating, in the bombing of Hiroshima, Japan, hundreds of thousands of civilians were killed as the result, 69% of buildings were destroyed, and 7% were damaged severely.

Overall the effects of nuclear

Table H2

*III Rubric Evaluation of Lower Scoring Essay*

Criterion	Score 0 = no 1 = somewhat, or one example 2 = yes definitely, or more than one example	Evidence/Justification
Does the persuasive essay make an argument consistent with the expectations outlined in the topic prompt.	0	No. No argument presented.
Does the persuasive essay include information learned from more than one source?	2	Yes. Blogspot & Wikipedia (Hiroshima)
Does the persuasive essay include information learned from more than one medium?	2	Yes Text & Video

Table H2 (cont'd)

In the persuasive essay, is the central argument/position grounded in corroborating facts from two or more websites/texts?	0	No. The information presented is from only 1 website. No evidence of corroboration across texts.
Does the persuasive argument include counterpoints to the central argument collected from one or more sources different from the sources used to construct the central argument?	0	No.
Does the persuasive essay integrate facts that were recorded as part of the author's bank of pre-existing knowledge?	2	Yes.
Does the essay provide evidence for construction of an integrated mental model of understanding: Is there evidence of integration of information across texts and/or within texts, and/or with background knowledge?	1	There is some evidence for integration of information within texts and with background knowledge, but not really across texts.
Does the persuasive essay include linguistic markers indicative of integration (e.g., seriation, transitional phrases that connect ideas, connectives, parallel structures that show an integrated understanding)	1	Yes -- She uses although "Nuclear power plants take up a lot of space and are very expensive to build, although once built they are inexpensive to maintain" --this shows intratextual integration because these ideas come from the same source. There are no transitional phrases.
Does the persuasive essay include explicit reference to source information [i.e. mention of author, a reason for why we should trust this information]?	0	No.
Does the persuasive essay include a thesis/synthesis statement that communicates an integrated understanding of the topic?	0	No.
	TOTAL:	8

## Appendix I: John and Alex's Evaluations

*Note:* Essays and annotations for John and Alex's work is provided in Section 3 of results, in text.

Table II

*John's Pretest TII Evaluation [Score = 1/20]*

Criterion	Score 0 = no 1 = somewhat, or one example 2 = yes definitely, or more than one example	Evidence/Justification
Does the persuasive essay make an argument consistent with the expectations outlined in the topic prompt.	0	No. It does not. There is no stance taken, other than personal preference. There is no clear thesis. There is an acknowledgement of the audience, but the focus is largely on what HU would choose for himself.
Does the persuasive essay include information learned from more than one source?	0	No. It includes information from only 1 source.
Does the persuasive essay include information learned from more than one medium?	0	No. Only text.
In the persuasive essay, is the central argument/position grounded in corroborating facts from two or more websites/texts?	0	No.

Table II (cont'd)

Does the persuasive argument include counterpoints to the central argument collected from one or more sources different from the sources used to construct the central argument?	0	No.
Does the persuasive essay integrate facts that were recorded as part of the author's bank of pre-existing knowledge?	0	No bkg recorded.
Does the essay provide evidence for construction of an integrated mental model of understanding: Is there evidence of integration of information across texts and/or within texts, and/or with background knowledge?	0	No evidence of integration across multiple texts. The information provided is sparse and traceable to only one of the two sources that they read closely.
Does the persuasive essay include linguistic markers indicative of integration (e.g., seriation, transitional phrases that connect ideas, connectives, parallel structures that show an integrated understanding)	1	He does use Although...and So...to juxtapose ideas in the second paragraph.
Does the persuasive essay include explicit reference to source information [i.e. mention of author, a reason for why we should trust this information]?	0	No.
Does the persuasive essay include a thesis/synthesis statement that communicates an integrated understanding of the topic?	0	No.
	TOTAL:	1

Table I2

*Alex's Pretest TII Evaluation [Score = 3/20]*

Criterion	Score 0 = no 1 = somewhat, or one example 2 = yes definitely, or more than one example	Evidence/Justification
Does the persuasive essay make an argument consistent with the expectations outlined in the topic prompt.	0	No. There is no argument made.
Does the persuasive essay include information learned from more than one source?	1	Two sources.
Does the persuasive essay include information learned from more than one medium?	0	No. Only txt.
In the persuasive essay, is the central argument/position grounded in corroborating facts from two or more websites/texts?	0	No corroboration.
Does the persuasive argument include counterpoints to the central argument collected from one or more sources different from the sources used to construct the central argument?	0	No counterpoints.
Does the persuasive essay integrate facts that were recorded as part of the author's bank of pre-existing knowledge?	0	No.
Does the essay provide evidence for construction of an integrated mental model of understanding: Is there evidence of integration of information across texts and/or within texts, and/or with background knowledge?	1	Evidence of integration within texts, and perhaps he recognized the way that he could juxtapose information from the two sources he uses to construct this essay -- but there is no integration of information across texts.

Table I2 (cont'd)

Does the persuasive essay include linguistic markers indicative of integration (e.g., seriation, transitional phrases that connect ideas, connectives, parallel structures that show an integrated understanding)	1	<p>I think these two sentences -- the end of one paragraph and the start of another show an attempt to integrate</p> <p>Doctors can give you advice on how to help these side effects.</p> <p>there is certain tasks you can do to help prevent side effects of treatment</p>
Does the persuasive essay include explicit reference to source information [i.e. mention of author, a reason for why we should trust this information]?	0	No
Does the persuasive essay include a thesis/synthesis statement that communicates an integrated understanding of the topic?	0	No.
	TOTAL:	3

Table I3

*John's Posttest TII Evaluation [Score = 7/20]*

Criterion	Score 0 = no 1 = somewhat, or one example 2 = yes definitely, or more than one example	Evidence/Justification
Does the persuasive essay make an argument consistent with the expectations outlined in the topic prompt.	1	Somewhat. He understands the general purpose of the prompt but makes no mention of parents as the intended audience. His position on the issue isn't actually based on the facts he reports -- but rather on his personal preference -- which is for a less expensive soap.
Does the persuasive essay include information learned from more than one source?	2	Yes. He provides facts collected from 3 sites.
Does the persuasive essay include information learned from more than one medium?	0	Only text.
In the persuasive essay, is the central argument/position grounded in corroborating facts from two or more websites/texts?	0	No. His central argument is based on information from ONE source -- the charearl.com site; a position that doesn't take into account any of the other facts he collected.
Does the persuasive argument include counterpoints to the central argument collected from one or more sources different from the sources used to construct the central argument?	1	Yes. They did search specifically for harmful effects of synthetic soaps on babies and they used the dherbs.com site to record information about the damaging effects of synthetic soap. He does include this info in the essay as a counterpoint to the "convenience/affordability" position.
Does the persuasive essay integrate facts that were recorded as part of the author's bank of pre-existing knowledge?	1	Yes. They discussed what synthetic meant and recorded their definition (manufactured/plastic)

Table I3 (cont'd)

Does the essay provide evidence for construction of an integrated mental model of understanding: Is there evidence of integration of information across texts and/or within texts, and/or with background knowledge?	1	<p>There is some evidence but not enough for a score of 2. He juxtaposes the negatives of synthetics with a positive quality of naturals: In my research i had found that thethe ingredients in some mass brands contain sodium laurel sulfate which causes eye damage in cases where the soap comes in contact with the eye. On another side of the spectrum soap made with nuatural products wash off cleanly where synthectic soaps leave a residue.</p> <p>But that's where the mutiple text inegration ends.</p>
Does the persuasive essay include linguistic markers indicative of integration (e.g., seriation, transitional phrases that connect ideas, connectives, parallel structures that show an integrated understanding)	1	e.g., On the other side of the spectrum
Does the persuasive essay include explicit reference to source information [i.e. mention of author, a reason for why we should trust this information]?	0	No.
Does the persuasive essay include a thesis/synthesis statement that communicates an integrated understanding of the topic?	0	No. It communicates an opinion but not an opinion informed by evidence from multiple sources.
	TOTAL:	7



Table I4

*Alex's Posttest TII Evaluation [Score = 10/20]*

Criterion	Score 0 = no 1 = somewhat, or one example 2 = yes definitely, or more than one example	Evidence/Justification
Does the persuasive essay make an argument consistent with the expectations outlined in the topic prompt.	1	His argument -- stated at the end is to use synthetics. However, he begins by listing the negatives of synthetics relative to natural soaps. He says naturals are safe -- but at the end he says that synthetics aren't really that bad, using his own experiences as proof. He does not maintain a consistent stance in this essay.
Does the persuasive essay include information learned from more than one source?	2	Yes.
Does the persuasive essay include information learned from more than one medium?	0	Only text.
In the persuasive essay, is the central argument/position grounded in corroborating facts from two or more websites/texts?	1	No. Actually. He uses as his main argument to use synthetic soaps their relative price and availability.
Does the persuasive argument include counterpoints to the central argument collected from one or more sources different from the sources used to construct the central argument?	1	Yes. They did search specifically for harmful effects of synthetic soaps on babies and they used the dherbs.com site to record information about the damaging effects of synthetic soap. He does include this info in the essay as a counterpoint to the "convenience/affordability" position.

Table I4 (cont'd)

Does the persuasive essay integrate facts that were recorded as part of the author's bank of pre-existing knowledge?	1	yes. He recorded his understanding of synthetic (manufactured) and has used this to guide his thinking in the essay.
Does the essay provide evidence for construction of an integrated mental model of understanding: Is there evidence of integration of information across texts and/or within texts, and/or with background knowledge?	2	yes -- there is evidence of integration across texts. e.g., "Natural soaps are safe for children and adults but they have there drawbacks too and  Natural soaps cost more and can be relitively hard to find and may have to be orderd. Mass produced soap is easy to find and cheap
Does the persuasive essay include linguistic markers indicative of integration (e.g., seriation, transitional phrases that connect ideas, connectives, parallel structures that show an integrated understanding)	1	He generally lists facts through this essay without connecting them The organizational sequence of ideas provides some coherence. He uses "but" as a way to juxtapose ideas from different texts. e.g. Natural soaps are safe for children and adults but they have there drawbacks too.
Does the persuasive essay include explicit reference to source information [i.e. mention of author, a reason for why we should trust this information]?	0	No.
Does the persuasive essay include a thesis/synthesis statement that communicates an integrated understanding of the topic?	1	There is a statement at the end, but his position is a waffly one.
	TOTAL:	10

## REFERENCES

## REFERENCES

- ACT (2006). *Persuasive Essay Analytic Rubric*. Iowa City, IA: ACT, Inc. Retrieved from [http://www.michigan.gov/mde/0,4615,7-140-38924\\_41644\\_42674---,00.html](http://www.michigan.gov/mde/0,4615,7-140-38924_41644_42674---,00.html)
- Afflerbach, P., & Cho, B.-Y. (2009). Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 69-90). New York: Routledge.
- Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (Vol. 1, pp. 255-291). New York: Longman.
- Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology*, 96(3), 523–535. doi:10.1037/0022-0663.96.3.523
- Azevedo, R. & Witherspoon, A.M. (2009). Self-regulated learning with hypermedia. In D.J. Hacker, J. Dunlosky, & A. Graesser (Eds.). *Handbook of metacognition in education*, (pp. 319-339). New York: Routledge.
- Bakeman, R. & Quera, V. (2013). GSEQ (Version 5.1.15) [Computer Software]. Retrieved from <http://www2.gsu.edu/~psyrab/gseq/Download.html>
- Bakeman, R. & Quera, V. (2011). *Sequential analysis and observational methods for the behavioral sciences*. New York: Cambridge University Press.
- Bakhtin, M.M. (1981). *The dialogic imagination: Four essays of M.M. Bakhtin*. M. Holquist (Ed.) (C. Emerson & M. Holquist, Trans.) Austin, TX: University of Texas Press.
- Balcytienne, A. (1999). Exploring individual processes of knowledge construction with hypertext. *Instructional Science*, 27, 303–328.
- Bazeley, P. (2013). *Qualitative data analysis: Practical strategies*. New York: Sage
- Bazerman, C. (1985). Physicists reading physics: Schema-laden purposes and purpose-laden schema. *Written Communication*, 2(1), 3-23.
- Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 47(3-4), 296–312. doi:10.1111/j.1469-7610.2006.01611.x
- Braasch, J. L. G., Lawless, K. A., Goldman, S. R., Manning, F. H., Gomez, K. W., & MacLeod, S. M. (2009). Evaluating search results: An empirical analysis of middle school students'

- use of source attributes to select useful sources. *Journal of Educational Computing Research*, 41(1), 63–82. doi:10.2190/EC.41.1.c
- Braasch, J. L. G. G., Bråten, I., Strømsø, H. I., Anmarkrud, Ø., & Ferguson, L. E. (2013). Promoting secondary school students' evaluation of source features of multiple documents. *Contemporary Educational Psychology*, 38(3), 180–195. doi:10.1016/j.cedpsych.2013.03.003
- Bråten, I., & Strømsø, H. I. (2011). Measuring strategic processing when students read multiple texts. *Metacognition and Learning*, 6(2), 111–130. doi:10.1007/s11409-011-9075-7
- 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(1), 6-28. doi:10.1598/RRQ.41.1.1
- Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, 20(4), 485-522.
- Britt, M.A., Rouet, J.-F. & Braasch, J. (2013). Documents as entities: Extending the situation model theory of comprehension. In M.A. Britt, S. R. Goldman & J.-F. Rouet (Eds.) *Reading-from words to multiple texts* (pp. 160-179). New York: Routledge.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.
- Castek, J. (2008). *How do 4th and 5th grade students acquire the new literacies of online reading comprehension? Exploring the contexts that facilitate learning.* (Doctoral dissertation). Retrieved from ProQuest dissertations and theses (3340875).
- Center for Educational Performance and Information (2013). *Free and reduced lunch counts.* Retrieved from: [http://www.michigan.gov/cepi/0,1607,7-113-21423\\_30451\\_36965---,00.html](http://www.michigan.gov/cepi/0,1607,7-113-21423_30451_36965---,00.html)
- Cerdán, R., & Vidal-Abarca, E. (2008). The effects of tasks on integrating information from multiple documents. *Journal of Educational Psychology*, 100(1), 209–222. doi:10.1037/0022-0663.100.1.209
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Coiro, J. (2011a). Predicting reading comprehension on the Internet: Contributions of offline reading skills, online reading skills, and prior knowledge. *Journal of Literacy Research*, 43(4), 352-392. doi:10.1177/1086296X11421979
- Coiro, J. (2011b). Talking about reading as thinking: Modeling the hidden complexities of online reading comprehension. *Theory Into Practice*, 50(2), 107-115. doi:10.1080/00405841.2011.558435

- Coiro, J., & Dobler, E. (2007). Exploring the online reading comprehension strategies used by sixth-grade skilled readers to search for and locate information on the Internet. *Reading Research Quarterly*, 42(2), 214-257. doi:10.1598/RRQ.42.2.2
- Coiro, J., Castek, J., & Guzniczak, L. (2011). Uncovering online reading comprehension processes: Two adolescents reading independently and collaboratively on the Internet. In P.J. Dunston et al. (Eds.) *60th yearbook of the Literacy Research Association* (pp. 354-369). Oak Creek, WI: Literacy Research Association.
- Coiro, J. & Hagerman, M.S. (2013). *Digging deeper: Online reading comprehension and collaborative inquiry*. Retrieved from <http://www.slideshare.net/mschirahagerman/7-1516-digging-deeperonlineinquiry-digiuri-24275493>
- Coiro, J., Knobel, M., Lankshear, C., & Leu, D. J. (2008). Central issues in new literacies and new literacies research. In J. Coiro, M. Knobel, C. Lankshear, & D. J. Leu (Eds.), *Handbook of research in new literacies* (pp. 1-21). New York: Lawrence Erlbaum Associates Taylor & Francis Group.
- Colwell, J., Hunt-Barron, S., & Reinking, D. (2013). Obstacles to developing digital literacy on the Internet in middle school science instruction. *Journal of Literacy Research*, 45(3), 295-324. doi: 10.1177/1086296X13493273
- Corbin, J. (2009). Taking an analytic journey. In J.M. Morse, P.N Stern, J. Corbin, B. Bowers, K. Charmaz & A.E. Clark (Eds.), *Developing grounded theory: The second generation* (pp. 35-53). Walnut Creek, CA: Left Coast Press, Inc.
- Coté, N., & Goldman, S. R. (1999). Building representations of informational text: Evidence from children's think aloud protocols. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 169-194). Mahwah, NJ: Lawrence Erlbaum Associates.
- Coté, N., & Goldman, S. R. (2004). Building representations of informational text: Evidence from children's think-aloud protocols. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (pp. 660-683). Newark, DE: International Reading Association.
- Coté, N., Goldman, S.R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes*, 25(1), 1-53. doi: 10.1080/01638539809545019
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed-methods approaches* (3rd ed.). Thousand Oaks, CA: Sage Publications Inc.
- Denscombe, M. (2003) *The good research guide for small-scale social research projects* (2<sup>nd</sup> ed.). Maidenhead, UK: Open University Press.

- DeSchryver, M.D. (2012). Toward a theory of web-mediated knowledge synthesis: How advanced learners used the web to construct knowledge about climate change behavior (Doctoral dissertation, Michigan State University) Retrieved from [http://www.academia.edu/4358786/Toward\\_a\\_Theory\\_of\\_Web-Mediated\\_Knowledge\\_Synthesis\\_How\\_Advanced\\_Learners\\_Used\\_the\\_Web\\_to\\_Construct\\_Knowledge\\_about\\_Climate\\_Change\\_Behavior](http://www.academia.edu/4358786/Toward_a_Theory_of_Web-Mediated_Knowledge_Synthesis_How_Advanced_Learners_Used_the_Web_to_Construct_Knowledge_about_Climate_Change_Behavior)
- Dirks, K.T., (1999). The effects of interpersonal trust on work group performance. *Journal of Applied Psychology*, 84, 445-455.
- Drew, S. V. (2013). Open up the ceiling on the common core state standards: Preparing students for 21st century literacy now. *Journal of Adolescent & Adult Literacy*, 56(4), 321–330. doi:10.1002/JAAL.00145
- Duke, N. K., & Pearson, P. D. (2002). Effective practices for developing reading comprehension. In A. E. Farstrup & S.J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 205–242). Newark, DE. International Reading Association.
- Duke, N.K., Pearson, P.D., Strachan, S. L., & Billman, A.K. (2011). Essential elements of fostering and teaching reading comprehension. In S.J. Samuels & A.E. Farstrup (Eds.), *What research has to say about reading instruction* (4th ed., pp. 51-93). Newark, DE: International Reading Association.
- Dwyer, B. (2010). *Scaffolding Internet reading: A study of a disadvantaged school community in Ireland*. (Doctoral dissertation, University of Nottingham). Retrieved from <http://etheses.nottingham.ac.uk/2426/>
- Field, A. (2009). *Discovering statistics using SPSS* (3<sup>rd</sup> Ed.). London, UK: Sage.
- Fox, E. & Alexander, P. A. (2009). Text comprehension: A retrospective, perspective and prospective. In S.E. Israel & G.G. Duffy (Eds.) *Handbook of research on reading comprehension* (pp. 227-239). New York: Routledge.
- Geidd, J. N. (2004). Structural magnetic resonance imaging of the adolescent brain. *Annals of the New York Academy of Sciences*, 1021, 77–85. doi: 10.1196/annals.1308.009
- Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C, Nugent, T.F., Herman, D.H., Clasen, L.S., Toga, A.W., Rapoport, J.L., Thompson, P.M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, 101(21), 8174–9. doi:10.1073/pnas.0402680101
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. Chicago, IL: Aldine.
- Goldman, S. R., Braasch, J. L. G., Wiley, J., Graesser, A. C., & Brodowinska, K. (2012).

- Comprehending and learning from Internet sources□: Processing patterns of better and poorer learners. *Reading Research Quarterly*, 47(4), 356–381. doi:10.1002/RRQ.027
- Goldman, S. R., & Rakestraw, J. A. (2000). Structural aspects of constructing meaning from text. In M. L. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 311–335). Mahwah, NJ: Lawrence Erlbaum Associates.
- Goldman, S. R., & Scardamalia, M. (2013). Managing, understanding, applying, and creating knowledge in the information age: Next-generation challenges and opportunities. *Cognition and Instruction*, 31(2), 255–269. doi:10.1080/10824669.2013.773217
- Goldman, S. R., & Wiley, J. (2004). Discourse analysis: Written text. In N. K. Duke & M. H. Mallette (Eds.), *Literacy research methods* (pp. 62-91). New York: Guilford Press.
- Guthrie, J. T., Mcrae, A., & Klauda, S. L. (2007). Contributions of concept-oriented reading instruction to knowledge about interventions for motivations in reading. *Educational Psychologist*, 4(4), 237–250. doi:10.1080/00461520701621087
- Guthrie, J. T., Van Meter, P., McCann, A. D., Wigfield, A., Bennett, L., Poundstone, C. C., Rice, M. E., et al. (1996). Growth of literacy engagement: Changes in motivations and strategies during concept-oriented reading instruction. *Reading Research Quarterly*, 31(3), 306–332.
- Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 403-424). Mahwah, NJ: Lawrence Erlbaum Associates.
- Guthrie, J. T., Wigfield, A., & Klauda, S. L. (2012). *Adolescents' engagement in academic literacy* (Report No. 7). Retrieved from [www.corilearning.com/research-publications](http://www.corilearning.com/research-publications)
- Hartman, D. K. (1995). Eight readers reading: The intertextual links of proficient readers reading multiple passages. *Reading Research Quarterly*, 30(3), 520-561.
- Hartman, D. K., Morsink, P. M., & Zheng, J. (2010). From print to pixels: The evolution of cognitive conceptions of reading comprehension. In E.A. Baker (Ed.) *The new literacies: Multiple perspectives and practice* (pp. 131-164). New York: Guilford Press.
- Henry, L. A. (2006). SEARCHing for an answer: The critical role of new literacies while reading on the Internet. *The Reading Teacher*, 59(7), 614-627.
- Hiebert, E. H., & Mesmer, H. A. E. (2013). Upping the ante of text complexity in the common core state standards: Examining its potential impact on young readers. *Educational Researcher*, 42(1), 44–51. doi:10.3102/0013189X12459802
- Hirsh, S. G. (1999). Children's relevance criteria and information seeking on electronic resources. *Journal of the American Society for Information Science*, 50(14), 1265–1283. doi:10.1002/(SICI)1097-4571(1999)50:14<1265::AID-ASI2>3.0.CO;2-E



- IBM (2013). IBM SPSS Statistics (v. 21). [Computer software]. Retrieved from <http://www-01.ibm.com/software/analytics/spss/products/statistics/downloads.html>
- Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of Research in Education*, 32, 241-267. doi: 10.3102/0091732X07310586
- Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York, NY: The Guilford Press.
- Kiili, C. (2013). Argument graph as a tool for promoting collaborative online reading. *Journal of Computer Assisted Learning*, 29(3), 248–259. doi:10.1111/j.1365-2729.2012.00492.x
- Kiili, C., Laurinen, L., Marttunen, M., & Leu, D. J. (2012). Working on understanding during collaborative online reading. *Journal of Literacy Research*, 44(4), 448-443. doi: 10.1177/1086296X12457166
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363-394. doi:10.1037/0033-295X.85.5.363
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 4(2), 75–86.
- Klein, P. D. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Educational Psychology Review*, 11(3), 203–271. doi:10.1023/A:1021913217147
- Klein, P. D., & Rose, M. A. (2010). Teaching argument and explanation to prepare junior students for writing to learn. *Reading Research Quarterly*, 45(4), 433–461. doi: 10.1598/RRQ.45.4.4
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. New York: Routledge.
- Kress, G., Jewitt, C., Ogborn, J., & Tsatsarelis, C. (2001). *Multimodal teaching and learning: The rhetorics of the science classroom*. London: Continuum International.
- Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. New York: Oxford University Press.
- Kucan, L., & Beck, I. L. (1997). Thinking aloud and reading comprehension research: Inquiry, instruction, and social interaction. *Review of Educational Research*, 67(3), 271-299. doi:10.3102/00346543067003271

- Kuhn, D. (2006). Do cognitive changes accompany developments in the adolescent brain? *Perspectives on Psychological Science*, 1(1), 59–67. doi:10.1111/j.1745-6924.2006.t01-2-.x
- Kymes, A. (2005). Teaching online comprehension strategies using think-alouds. *Journal of Adolescent and Adult Literacy*, 48(6), 492-500.
- Landis, J.R. & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.
- Langer, J. A. (1986a). Learning through writing□: Study skills in the content areas. *Journal of Reading*, 29(5), 400–406.
- Langer, J. A. (1986b). Reading, writing, and understanding: An analysis of the construction of meaning. *Written Communication*, 3(2), 219–267. doi:10.1177/0741088386003002005
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.
- Leu, D. J. (2000). Literacy and technology: Deictic consequences for literacy education in an information age. In M. L. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 743–770). Mahwah, NJ: Lawrence Erlbaum Associates.
- Leu, D.J., Coiro, J., Castek, J., Hartman, D.K., Henry, L., Reinking, D. (2008). Research on instruction and assessment in the new literacies of online reading comprehension. In C.C. Block & S. R. Parris (Eds.) *Comprehension instruction: Research-based best practices* (2nd ed., pp. 321-346). New York: Guilford Press.
- Leu, D. J., Coiro, J., Kulikowich, J., Sedransk, N., Everett-Cacopardo, H., McVerry, G., O’Byrne, W. I., et al. (April, 2012). *Developing three formats for assessing online reading comprehension: The ORCA project year 3*. Paper session presented at the meeting of the American Educational Research Association. Vancouver, Canada.
- Leu, D. J., Kinzer, C. K., Coiro, J., & Cammack, D. W. (2004). Toward a theory of new literacies emerging from the Internet and other information communication technologies. In R. Ruddell & N. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1568-1611). Newark, DE: International Reading Association.
- Leu, D.J. & Reinking, D. (2005a) *Teaching internet comprehension to adolescents (TICA) project*. Retrieved from <http://www.newliteracies.uconn.edu/iesproject/>
- Leu, D.J. & Reinking, D. (2005b) *TICA overview*. Retrieved from <http://www.newliteracies.uconn.edu/iesproject/documents/IRTOverview.pdf>

- Lindgaard, G., Dudek, C., Sen, D., Sumegi, L., & Noonan, P. (2011). An exploration of relations between visual appeal, trustworthiness and perceived usability of homepages. *ACM Transactions on Computer-Human Interaction*, 18(1), 1–30. doi:10.1145/1959022.1959023
- Macedo-Rouet, M., Braasch, J. L. G., Britt, M. A., & Rouet, J. (2013). Teaching fourth and fifth graders to evaluate information sources during text comprehension. *Cognition and Instruction*, 31(2), 204-226. doi:10.1080/07370008.2013.769995
- Mateos, M., Martín, E., Villalón, R., & Luna, M. (2008). Reading and writing to learn in secondary education: Online processing activity and written products in summarizing and synthesizing tasks. *Reading and Writing*, 21, 675-697. doi:10.1007/s11145-007-9086-6
- Mayer, R. E. (1979). Can advance organizers influence meaningful learning? *Review of Educational Research*, 49(2), 371. doi:10.2307/1169964
- Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *The American Psychologist*, 59(1), 14–9. doi:10.1037/0003-066X.59.1.14
- McNamara, D.S. & Shapiro, A.M. (2005). Multimedia and hypermedia solutions for promoting metacognitive engagement, coherence, and learning. *Journal of Educational Computing Research*, 33(1), 1-29. doi: 10.2190/7N6R-PCJL-UMHK-RYPJ
- Michigan Department of Education. (2006). *Essential high school science content expectations*. East Lansing, MI: Michigan Department of Education. Retrieved from [http://www.michigan.gov/mde/0,1607,7-140-38924\\_41644\\_42814---,00.html](http://www.michigan.gov/mde/0,1607,7-140-38924_41644_42814---,00.html)
- Miles, M.B. & Huberman, A.M. (1994). *Qualitative data analysis: A sourcebook of new methods*. Thousand Oaks, CA: Sage.
- Moje, E.B., Stockdill, D., Kim, K. & Kim, H-J. (2011). The role of text in disciplinary learning. In M.L. Kamil, P.D. Pearson, E.B. Moje & P.P. Afflerbach (Eds.) *Handbook of reading research* (Vol. 4, pp. 453-486). New York: Routledge.
- National Governors Association Center for Best Practices & The Council of Chief State School Officers (2010). *Common core state standards for English language arts & literacy in history/social studies, science and technical subjects*. Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers. Retrieved from <http://www.corestandards.org/the-standards>
- Negretti, R. (2012). Metacognition in student academic writing: A longitudinal study of metacognitive awareness and its relation to task perception, self-regulation, and evaluation of performance. *Written Communication*, 29,142–179. doi:10.1177/0741088312438529

- Newell, G.E. (2006). Writing to learn: How alternative theories of school writing account for student performance. In C.A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 235-247). New York: Guilford.
- Newell, A. & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. *Cognition and Instruction*, 2, 117-175.
- Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, 8, 317-344.
- Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Toward 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.
- Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA. Retrieved from [http://www.rand.org/pubs/monograph\\_reports/MR1465.html](http://www.rand.org/pubs/monograph_reports/MR1465.html)
- Reinking, D., & Bradley, B. A. (2008). *Formative and design experiments: Approaches to language and literacy research*. New York: Teachers College Press and National Conference on Research in Language and Literacy (NCRL).
- Roswell, J., & Burke, A. (2009). Reading by design: Two case studies of digital reading practices. *Journal of Adolescent & Adult Literacy*, 53(2), 106-118.  
doi:10:1598/JAAL.53.2.2
- Rouet, J.-F., Favart, M., Britt, M. A., & Perfetti, C. A. (1997). Studying and using multiple documents in history□: Effects of discipline expertise. *Cognition and Instruction*, 15(1), 85-106.
- Rouet, J.-F. (2006). *The skills of document use: From text comprehension to web-based learning*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Sevensma, K. (2013). *Negotiating new literacies in science: An examination of at-risk and average-achieving ninth-grade readers' online reading comprehension strategies*. (Doctoral dissertation). Retrieved from ProQuest, UMI Dissertations Publishing (2013. 3563033.)
- Shanahan, C. (2009). Disciplinary comprehension. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 240-260). New York: Routledge.

- Small, G. W., Moody, T. D., Siddarth, P., & Bookheimer, S. Y. (2009). Your brain on Google: Patterns of cerebral activation during Internet searching. *American Journal of Geriatric Psychiatry*, 17(2), 116–126.
- Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (2004). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. In R. B. Ruddell & N. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 640-653). Newark, DE: International Reading Association.
- Spiro, R., & DeSchryver, M. (2009). Constructivism: When it's the wrong idea and when it's the only idea. In S. Tobias & T. M. Duffy (Eds.), *Constructivist instruction: Success or failure?* (pp. 106–123). New York: Routledge.
- Spivey, N. N., & King, J. R. (1989). Readers as writers composing from sources. *Reading Research Quarterly*, 24(1), 7-26.
- Stahl, S. A., Hynd, C. R., Britton, B. K., McNish, M. M., & Bosquet, D. (1996). What happens when students read multiple source documents in history? *Reading Research Quarterly*, 31(4), 430-456.
- Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends in Cognitive Sciences*, 9(2), 69–74. doi:10.1016/j.tics.2004.12.005
- Strømsø, H. I., Bråten, I., Samuelstuen, M. S. (2003). Students' strategic use of multiple sources during expository text reading□: A longitudinal think-aloud study. *Cognition and Instruction*, 21(2), 113-147.
- Techsmith Corporation (2012). Camtasia for Mac (Version 2.2.2). [Computer software]. East Lansing, MI: Techsmith.
- Techsmith Corporation (2012). Morae (Version 3.3.2). [Computer software]. East Lansing, MI: Techsmith.
- United States Census Bureau (2012). *State and county quickfacts*. Retrieved from <http://quickfacts.census.gov/qfd/index.html>
- Vaughn, S., Klingner, J. K., Swanson, E. A., Boardman, A. G., Roberts, G., Mohammed, S. S., & Stillman-Spisak, S. J. (2011). Efficacy of collaborative strategic reading with middle school students. *American Educational Research Journal*, 48(4), 938-964. doi:10.3102/0002831211410305
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

- Wang, Y. D., & Emurian, H. H. (2005). An overview of online trust: Concepts, elements, and implications. *Computers in Human Behavior*, 21(1), 105–125. doi:10.1016/j.chb.2003.11.008
- 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.
- Wiley, J., Goldman, S.R., Graesser, A.C., Sanchez, C.A., Ash, I.K., & Hemmerich, J.A. (2009). Source evaluation, comprehension and learning in Internet science inquiry tasks. *American Educational Research Journal*, 46(4), 1060-1106. doi: 10.3102/0002831209333183
- Wineburg, S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, 83, 73-87.
- Wolfe, M. B. W., & Goldman, S. R. (2005). Relations between adolescents' text processing and reasoning. *Cognition and Instruction*, 23(4), 467-502. doi:10.1207/s1532690xc2304\_2
- Woodcock, R.W. (2011). *Woodcock Reading Mastery Tests* (3<sup>rd</sup> Ed.). San Antonio, TX: Pearson Education.
- Yin, R. K. (1989). *Case study research: Design and methods* (Revised.). Newbury Park, CA: Sage Publications, Inc.
- Zhang, S., & Duke, N. K. (2008). Strategies for Internet reading with different reading purposes: A descriptive study of twelve good Internet readers. *Journal of Literacy Research*, 40, 128–162.
- Zhang, S., & Duke, N. K. (2011). The impact of instruction in the WWWDOT framework on students' disposition and ability to evaluate web sites as sources of information. *The Elementary School Journal*, 112(1), 132–154. doi:10.1086/660687