

ABSTRACT

Title of Document: MODELING MULTIPLE SOURCE USE: USING LEARNER CHARACTERISTICS AND SOURCE USE BEHAVIORS TO PREDICT RESPONSE QUALITY

Alexandra List, Doctor of Philosophy, 2014

Directed By: Professor Patricia A. Alexander
Human Development and Quantitative Methodology

Multiple source use (MSU) has been identified as both a critical competency and a key challenge for today's students, living in the digital age (Goldman & Scardamalia, 2013b). Theoretical models of multiple source use provide insights into how the MSU process unfolds and identify points at which students may encounter challenges (i.e., in source selection, processing, and evaluation, Rouet & Britt, 2006). However, understandings of MSU have been limited by two gaps in the literature. First, while points of challenge in students' MSU process have been examined independently, comprehensive models considering the joint role of source selection, processing, and evaluation in task performance have not been fully investigated. Further, while research on MSU has focused on students' behaviors when engaging with texts, individual difference factors have been considered only to a limited extent, despite their theorized importance (Rouet, 2006).

The purpose of the present study was to examine the extent to which multiple source use behaviors (i.e., source selection, processing, and evaluation) and learner characteristics (i.e., prior knowledge, domain general source evaluation behaviors, stances on the target issue) predicted open-ended task performance.

Participants were 197 undergraduate students, asked to complete measures assessing their prior knowledge, stances on the Arab Spring in Egypt, the topic of the task,

and domain general source evaluation behaviors. Then, participants were tasked with using a library of six sources to respond to a controversial prompt about a contemporary event (i.e., Arab Spring in Egypt). Four indices were used to assess open-ended response quality: (a) word count, (b) the number of arguments included in students' responses, (c) scores on the SOLO taxonomy (Biggs & Collis, 1982), reflecting the extent to which responses integrated and evaluated information presented across texts, and (d) the number of citations in students' answers.

Key findings included the role of students' ratings of source interestingness and time on texts as predictive of open-ended task performance. Further, students' trustworthiness evaluations were found to be associated with SOLO scores. Overall, as compared to multiple source use behaviors, learner characteristics were found to have a more limited effect on task performance.

MODELING MULTIPLE SOURCE USE: USING LEARNER CHARACTERISTICS
AND SOURCE USE BEHAVIORS TO PREDICT RESPONSE QUALITY

By

Alexandra List

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2014

Advisory Committee:

Professor Patricia A. Alexander, Chair
Professor Peter Afflerbach
Assistant Professor Liliana Maggioni
Professor Kathryn R. Wentzel
Professor Allan Wigfield

ACKNOWLEDGMENTS

In my first semester of graduate school I took Achievement Motivation with Dr. Kathryn Wentzel. In one lecture, we discussed the instrumental value of social support for academic achievement. At the time, this was something I had never thought about but reflecting on my graduate school career, no lesson has proven truer. I am truly privileged to have many people to thank and to be indebted to many, many more.

I want to start by thanking my parents and grandmother, who have worked so hard to afford me every possible opportunity. They continue to be the best examples I know of what it means to be caring, good people. I also want to thank my sister, Yana. In addition to being my best friend, she is a dedicated, reflective, and energetic teacher who is the embodiment of my faith in education.

My “chosen” family is no less exceptional. Molly Weston and Lynn Conell-Price, after a decade of friendship, continue to support me in all ways, both big and small. As planners, we talk a lot about the future, but looking at the present, I am amazed at where they are and where they’re going. Matthew Tilghman likely holds the distinction of being the person who has seen me cry the most. I am ever appreciative of his cool, levelheadedness and pragmatism. I want to especially acknowledge Cyrus Stoller. He is the smartest person I know. I love his passion for ideas and his enthusiasm for sharing projects. My dissertation would not have been possible without Cyrus’s talents and hard work in developing the data collection website. We have always been invested in each other’s projects and being able to collaborate with your best friend is a unique gift.

The family I was “assigned” came to me unexpectedly when I joined the Disciplined Reading and Learning Research Lab. I have been so fortunate to benefit for

the mentoring of my older and wiser siblings: Emily Fox, Liliana Maggioni, Dan Dinsmore, Megan Parkinson, Sandra Loughlin (who also doubles as my life coach), and Lily Fountain and to watch younger members, Sophie Jablansky, Lauren Singer, Courtney Hattan, Denis Dumas (technically the precise same age), and Peter Baggetta, acclimate and flourish. All of them have been role models for what it means to be an exceptional researcher and I sincerely hope they will continue to be good friends. Collaborating with Emily Grossnickle has been instrumental in my development as a graduate student and a scholar. From my first year when she advised me to get a parking pass, she has been the yin to my research yang. I am so excited about her dissertation, reflecting her creativity and diligence as a researcher, and for many more projects to come.

Very rightly so, the lab is called the “Alexander Family”. Patricia Alexander has been an all-encompassing force in my life for the past six years –and I would not have had it any other way. I admire so much about Patricia – her passion for educational psychology, her dedication to her students, her generosity of spirit, and most of all, the intellectual rigor with which she approaches the world. She is among the most remarkable people I know and I feel privileged, in addition to calling her my advisor, mentor, and toughest critic, to call her a great friend.

I want to thank my committee for all of the time and thought they devoted to this project. Certainly the document is much strengthened based on feedback from Peter Afflerbach, Liliana Maggioni, Kathy Wentzel, and Allan Wigfield. I am especially grateful for their insistence that the project be expanded. Having been a student for a long time, I was nervous to be done with courses and so pleased to have the fortune of

demanding, skeptical, insightful faculty still available for guidance. I know my committee to be exceptional thinkers and scholars as well as kind and caring mentors.

Finally, I want to acknowledge those people who were there when it all began. I had a fantastic cohort, including Lauren Musu-Gillette, Colleen Morrison, Amy Ho, Amanda Mason, and Ting Zhang, who are all doing amazing things in education. They have been the best of friends and commiserators, as have Merly Barofsky, Jenna Cambria, Melissa Duchene, Alice Donlan, Sara Eason, Danette Morrison, Katie Muenks, Alexis Williams, and many others. Thank you to everyone mentioned and not– I am truly blessed to have had all of you in my life.

Table of Contents

CHAPTER 1: STATEMENT OF THE PROBLEM	1
Theories of Multiple Source Use	4
MD-TRACE Model of Multiple Source Use.....	5
Individual Differences in Multiple Source Use	7
Prior knowledge.....	7
Domain general source evaluation behaviors.	8
Stance.....	10
Gaps in the Literature.....	14
The Current Investigation	15
Research Questions	19
Key Terms.....	20
CHAPTER 2: REVIEW OF THE LITERATURE.....	25
Model of Multiple Source Use.....	25
Step 1: Task Model Construction	27
Step 2: Determination of Information Need	27
Step 3: Source Processing	28
Source selection.	29
Source processing.	30
Source evaluation.....	32
Step 4: Response Formulation	32
Step 5: Response Evaluation.....	33
Internal and External Resources	33
Documents Model of Multiple Source Use	36
Sourcing, Source Evaluation, and the Documents Model	38
Processes of Multiple Source Use.....	40
Source Selection.....	40
Source Processing	46
Source Evaluation	54
Individual Difference Factors in Multiple Source Use	68
Prior Knowledge	69
Prior knowledge and source selection.....	70
Prior knowledge and text processing.	72
Selecting relevant information.....	74
Inference generation.....	74
Reduced cognitive load.....	75
Information integration.	75
Text organization.	76
Prior knowledge and source evaluation.	78
Stance on Target Topic	81
Domain General Source Evaluation Behaviors	87

Empirical Examinations of Models of Multiple Source Use	89
CHAPTER 3: METHODOLOGY	92
Participants.....	92
Sample Representativeness.....	93
Sample Size Estimation	95
Undergraduate Students as a Target Sample	97
Non-Expert Sample Status.....	99
Measures	100
Session 1: Learner Characteristics	100
Demographics.	100
Prior knowledge.	101
Stance.....	101
Domain general source evaluation behaviors.	103
Session 2: Multiple Source Use	105
Multiple source use task.....	105
Task topic.....	107
Task formulation.....	110
Texts.....	111
Text number.	114
Text length.	115
Text presentation.....	115
Source use data.	119
Accessing document information.	119
Source evaluation.....	120
Response composition.	123
Response coding.	123
Word count.....	123
Argument.	124
SOLO taxonomy.	124
Citation.....	127
Post task measures.	127
Data Analysis Plan.....	128
Multiple Regression	132
Regression vis-a-vis Path Analysis.....	134
CHAPTER 4: RESULTS	135
Descriptives of Multiple Source Use	137
Multiple Source Use Behaviors	137
Multiple Source Use Behaviors By Source Type	138
Task Performance	142
Task Performance by Source Access	144
Associations with Task Engagement	148
Multiple Regression Assumption Checking	150
Normality	152

Homoscedasticity and Linearity	153
Multicollinearity.....	155
Outliers.....	155
Multiple Source Use Behaviors Predicting Response Quality	158
Individual Difference Factors and MSU Behaviors.....	161
Learner Characteristics Predicting Response Quality.....	163
Learner Characteristics and Multiple Source Use Behaviors Predicting Response Quality.....	165
CHAPTER 5: DISCUSSION	170
Conclusions and Implications	171
Research Question 1	171
Research Question 2	175
Research Question 3	180
Research Question 4	182
Research Question 5	183
MD-TRACE Model	186
Conceptual Model.....	189
Limitations and Future Directions	195
Examining Strategies	198
Optimal Non-Expert Multiple Source Use	199
Expanded Analyses.....	202
Design Parameters	203
Generalizability.....	205
Implications for Practice.....	210
APPENDICES.....	212
Appendix A: Demographic Characteristics and Multiple Source Use	212
Appendix B: Demographics Questionnaire	221
Appendix C: Prior Knowledge Identification Measure	222
Appendix D: Stance Items	223
Appendix E: Credibility Assessment Scale	224
Appendix F: Academic Behaviors Questionnaire.....	225
Appendix G: Screenshot of Directions Page	227
Appendix H: Screenshot of Response Window.....	228
Appendix I: Texts in the Source Library	229
Appendix J: Screenshot of Source Library.....	235
Appendix K: Screenshots of Accessible Document Information	236

Appendix L: Screenshot of Source Rating Window.....	238
Appendix M: Sample Response Coding.....	239
Appendix N: Screenshot of Source Ranking and Justification Page.....	242
Appendix O: Post-Task Engagement Scale.....	243
Appendix P: Ascertaining Normality of Learner Characteristics and Multiple Source Use Behavioral Variables.....	244
Appendix Q: Residual P-P Plots for Predictive Models.....	248
Appendix R: Histograms of Residuals for Predictive Models.....	255
Appendix S: Partial Plots for Each Predictive Model.....	257
REFERENCES.....	265

List of Tables

Table 1. Demographic Comparison of Sample to Overall University Population	88
Table 2. Comparison of Traditional History Topics Used in Prior Research and the Contemporary Political Topic in the Current Investigation	102
Table 3. Document Information for Texts in the Multiple Source Use Task	103
Table 4. Variables Examined	120
Table 5. Descriptives of Variables Examined	127
Table 6. Descriptives of Students' Source Access	129
Table 7. Descriptives of MSU by Source Type	130
Table 8. Performance Descriptives by Source Type	136
Table 9. Correlation of Individual Difference Factors and Behavioral Variables and Task Engagement.....	139
Table 10. Correlation of Response Metrics and Task Engagement.....	140
Table 11. Correlation Matrix of Independent Variables	146
Table 12. Multicollinearity Statistics	146
Table 13. Multiple Source Use Behaviors Predicting Word Count.....	148
Table 14. Multiple Source Use Behaviors Predicting Number Of Arguments.....	149
Table 15. Multiple Source Use Behaviors Predicting Solo Scores.....	149
Table 16. Multiple Source Use Behaviors Predicting Number Of Citations	150
Table 17. Normality Statistics For Predictor Variables	151
Table 18. Spearman's Correlation Matrix for Individual Difference Factors and Source Use Behaviors	153
Table 19. Learner Characteristics Predicting Word Count.....	153
Table 20. Learner Characteristics Predicting Arguments	154
Table 21. Learner Characteristics and Multiple Source Use Behaviors Predicting Word Count.....	156
Table 22. Learner Characteristics and Multiple Source Use Behaviors Predicting Number Of Arguments	156
Table 23. Learner Characteristics And Multiple Source Use Behaviors Predicting SOLO Scores	157
Table 24. Learner Characteristics And Multiple Source Use Behaviors Predicting Number of Citations	158
Table 25. Summary of Predictions.....	168
Table 26. Differences in Learner Characteristics, Multiple Source Use, and Performance by Gender.....	198
Table 27. Differences in Multiple Source Use and Performance by Minority Status	199
Table 28. Differences in Multiple Source Use and Performance by Class Standing	200
Table 29. Differences in Multiple Source Use and Performance by Area of Study.....	201

Table 30. Demographics, Learner Characteristics, and Multiple Source Use Behaviors Predicting Word Count	202
Table 31. Demographics, Learner Characteristics, and Multiple Source Use Behaviors Predicting Number of Arguments	203
Table 32. Demographics, Learner Characteristics, and Multiple Source Use Behaviors Predicting SOLO Scores	204
Table 33. Demographics, Learner Characteristics, and Multiple Source Use Behaviors Predicting Citations	205
Table 34. Item-Specific Reliability Statistics	227

List of Figures

Figure 1. A graphic depiction of the Multiple Documents Task Based Relevance Assessment and Content Extraction model.....	6
Figure 2. Components of the conceptual model of individual difference factors and behaviors in multiple source use	13
Figure 3. Model of learner characteristics and multiple source use behaviors predicting open-ended response quality, examined in this study.....	18
Figure 4. Frequency of each number of sources accessed	128
Figure 5. Standardized time spent and source evaluation ratings by source type.....	132
Figure 6. Histograms of each open-ended outcome variable, word count, number of arguments, SOLO scores, and number of citations.....	134
Figure 7. Performance by type of source accessed	137
Figure 8. Access, document information use, and citation by source type.....	138

CHAPTER 1: STATEMENT OF THE PROBLEM

You got a logic in your house...it's got keys instead of dials and you punch the keys for what you wanna get...if you punch for the weather forecast or who won today's race at Hialeah or who was mistress of the White House during' Garfield's administration or what is PDQ and R sellin' for today, that comes on the screen too. The relays in the tank do it. The tank is a big buildin' full of all the facts in creation an' all the recorded telecasts that ever was made – an' it's hooked in with all the other tanks all over the country—an' anything you wanna know or see or hear, you punch for it an' you get it. Very convenient. Also it does math for you, an' keeps books, an' acts as a consultin' chemist, physicist, astronomer, an' tealeaf reader

The opening excerpt comes from the short story, *A Logic Named Joe*, written by science-fiction author Murray Leinster and published in the 1946 issue of *Astounding Science Fiction*. This story is widely considered to be the inspiration for the World Wide Web (Ferro & Swedin, 2009). The preceding excerpt opens this document because the ease of information access that Lienster foretold in 1946 is a reality for today's students. Although, as Lienster describes, today's "logics" provide students with equally easy access to information from "chemist, physicist, astronomer, an' tealeaf reader," they do not aid learners in distinguishing among these informational sources or in integrating their predictions into a viable response. Indeed, students, even at the undergraduate level, have been found to experience difficulties in effectively using their "logics." Specifically, students have been found to experience challenges in selecting from among

the multitude of varied sources their logics proffer, processing and making use of “all the facts in creation”, and in evaluating the credibility of “recorded telecasts” ranging widely in reliability (Britt & Aglinskias, 2002; Kiili, Laurinen, & Marttunen, 2008; Grimes & Boening, 2001; Metzger, Flanagin, & Zwarun, 2003). At the same time, these processes dominate students’ interactions with information online and are increasingly central to both academic and real-world success (Goldman & Scardamalia, 2013a; Lankshear & Knobel, 2011).

Many argue that the realization of what was science fiction only sixty years ago has come at a cost (e.g., Rothenberg, 1998). Since the introduction of the Internet and Web-based information sources into classrooms and curricula, there has been a correspondent outcry from educators and researchers about the questionable legitimacy of online sources and students’ limited abilities in evaluating texts (e.g., Grimes & Boening, 2001; Wang & Artero, 2005). Metzger (2007) outlines a number of reasons why such concerns may be well founded. For one, the Internet has no centralized standards for determining information quality and, unlike with traditional print texts, there are few gatekeepers to designate what may constitute an authoritative source. For another, unlike print sources, information on the Internet often does not undergo an editorial process (Flanagin & Metzger, 2000; Lorenzen, 2001). Further, online, information about author and publisher, features used to determine source reliability, may be difficult to locate (Britt & Gabrys, 2002) and unreliable sources may be made to appear credible with ease (Metzger, 2007). Perhaps more problematically, when conducting research online, learners have been found to focus on superficial factors when selecting and evaluating sources (e.g., relevance and layout) rather than considering factors related to source

credibility or information accuracy (Cottrell, 2001; Kammerer & Gerjets, 2012b).

Collectively the issues raised by Metzger (2007) and others (Smith, 1997; Johnson & Kaye, 1998; McInerney & Bird, 2005) suggest that using and critically evaluating sources on the Internet may be especially challenging for students at the same time that many of the sources that students are called to judge may be of a more questionable nature (Grimes & Boening, 2001; Metzger et al., 2003). Beyond presenting students with a great volume and variety of sources at ever-increasing speeds, the Internet has introduced students to entirely new types of texts as well (Coiro, 2003a). Indeed, in the last 10 years, Wikipedia entries, blogs, and tweets, previously unheard of, have begun to serve as common sources of information for today's learners (e.g., Brenner, 2012; Head & Eisenberg, 2010; Lim, 2009). Sources such as these, as well as other sites on the Internet, have been both praised for the broad and easy access to information they afford and critiqued for their questionable reliability and lack of editorship (Coiro, 2003b; Purcell et al., 2012). For instance, despite the availability and accessibility associated with sources like Wikipedia, students have been repeatedly cautioned to avoid such sites because of their seemingly non-scholarly qualities (Grimes & Boening, 2001; Head & Eisenberg, 2010).

Given students' use of new and controversial sources in their research process, such sources merit further investigation to understand how they are used, integrated, and evaluated by learners, particularly in relation to source types more typically found in academic contexts. Even beyond new source types, texts on the Internet have unique properties. Sources on the Internet vary widely in their formatting (Ciolek, 1996), emphasize currency, present an overwhelming volume of available information

(Eliopoulos & Gotlieb, 2003; Henry, 2006), and are commonly developed explicitly to, “sell, discredit, deceive, or persuade” (Coiro, 2003b, p. 29). These qualities suggest that students’ source selection, use, and evaluation may need to be specifically examined when learners encounter texts on the Internet. Further, Coiro’s (2003b) description of the Internet context as offering free, fast, and easy access to a great volume of information suggests reasons why selecting, judging, and integrating sources may be challenging for students who have to contend with an overwhelming amount of information before arriving at an understanding of an issue or completing an academic task.

Like Coiro (2003b), Mason, Boldrin, and Ariasi (2010a) conceptualize the Internet context as affording both benefits and challenges to today’s learners. While students are provided with easy and instantaneous access to a multitude of sources, they are, at the same time, required to develop complex skills associated with needing to continuously select among and evaluate the quality of these same sources in information saturated digital environments (Coiro, 2003b; Mason et al., 2010a). There has been a call for the development of a better understanding of how students select, use, and integrate sources in new and more precarious online contexts (Lorenzen, 2001; Walraven, Brand-Gruwel, & Boshuizen, 2009). Indeed, some researchers have gone so far as to recommend the examination of students’ source processing skills on the Internet as manifestations of *new* literacy skills altogether (Leu, Kinzer, Coiro, & Cammack, 2004).

Theories of Multiple Source Use

Rouet and Britt (2011) argue the need for a specific model to capture students’ processing when engaged in multiple source use, as opposed to single text comprehension, because multiple texts come to students in new forms and introduce new

uses. Specifically, Rouet and Britt (2011) propose the Multiple Documents Task-Based Relevance Assessment and Content Extraction Model (MD-TRACE) to capture the resources and processes that come into play when students engage in multiple text use. The MD-TRACE model conceptualizes students' interactions with multiple texts as occurring procedurally, in a series of iterative steps, including learners' source selections, use, and evaluations. Further, the MD-TRACE model suggests that students' interactions with multiple sources are guided by *internal resources* or individual level cognitive reserves that shape and support engagements with multiple texts.

MD-TRACE Model of Multiple Source Use

The MD-TRACE model outlines the multiple source use process as unfolding through five core phases or steps. In Step 1, learners construct a *task model*, or cognitive representation of task demands and how these may be satisfied (Rouet & Britt, 2011). The importance of the task model in driving students' subsequent interactions with multiple texts has been reinforced across studies examining differences in students' source use in response to task manipulations (e.g. Gil, Bråten, Vidal-Abaraca, & Strømsø, 2010a; 2010b; Wiley & Voss, 1999). In Step 2 of the MD-TRACE model, students make a determination of *information need*, or decide that they have insufficient information to meet task demands, thereby electing to access information and engage in multiple source use.

Step 3 of the model includes students' specific interactions with multiple texts. There are three sub-steps capturing students' text engagement: *source selection*, *source processing*, and *source integration*, predicated on *source evaluation*. These three sub-processes are considered to be cyclical. For instance, evaluating a source as unreliable

may lead students to select an alternate text or processing a text may lead students to judge it too difficult to comprehend. Once students conclude that their information needs have been met through multiple source use, they move to Step 4, or *formulate a response* to meet task demands. The fifth and concluding step of the MD-TRACE model involves students' determinations that their generated responses meet task demands and are in accordance with criteria set out in their initially developed task model (Step 1). A diagram of the MD-TRACE model is presented in Figure 1.

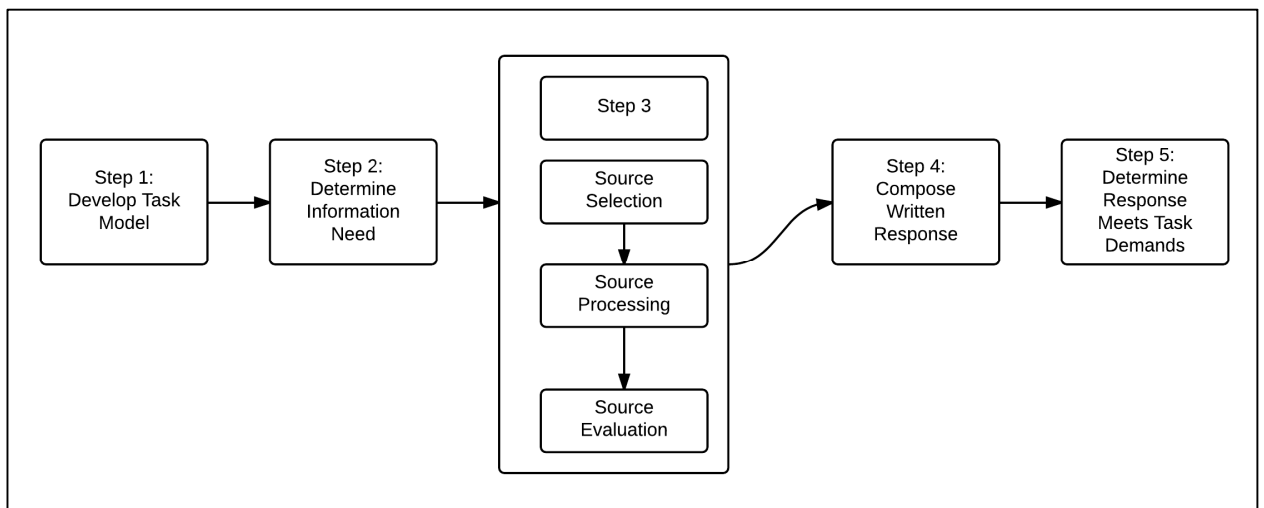


Figure 1. A graphic depiction of the Multiple Documents Task Based Relevance Assessment and Content Extraction model.

Empirical work has used the MD-TRACE model as a guiding framework to consider the specific steps or processes characteristic of students' multiple source interactions. In fact, students' source selections (Macedo-Rouet, Braasch, Britt, & Rouet, 2013), processing (McCrudden, Magliano, & Schraw, 2010), and evaluations (Bråten, Strømsø, & Britt, 2009; Bråten, Strømsø, & Salmerón, 2011) have all been examined in prior research, the latter receiving the most attention (e.g., Wiley et al., 2009). However, the literature has been limited in fully considering the MD-TRACE model by jointly examining students' source selections, processing, and evaluations. A primary goal of

the present study was to comprehensively examine the MD-TRACE model by exploring students' source selections, processing, and evaluations within a single model. The extent to which these behaviors predict the quality of students' written products, generated in response to a multiple source task was also investigated.

Individual Differences in Multiple Source Use

The MD-TRACE is a procedural model, focusing on the specific behaviors or cognitive processes in which learners engage when interacting with and integrating multiple texts. As in single text processing (Alexander, Kulikowich, & Jetton, 1994; Alexander, Kulikowich, & Schulze, 1994; Kintsch, 1994), students' multiple source use is impacted by a variety of individual difference factors, such as prior knowledge (Britt & Aglinskias, 2002; Rouet, Britt, Mason, & Perfetti, 1996; Wineburg, 1991). Within the MD-TRACE model, Rouet and Britt (2011) refer to such individual difference factors as *internal or cognitive* resources affecting students' multiple source use. These internal resources are considered to function in two primary ways. First, they shape the task model that students construct and second, they serve as cognitive reserves that students may draw on to aid in source use and task completion, particularly when encountering challenges (Rouet & Britt, 2011).

Prior knowledge. While the literature on multiple source use has been somewhat limited in examining individual difference factors, prior knowledge has often been investigated (Le Bigot & Rouet, 2007; Potelle & Rouet, 2003). Prior knowledge has been conceptualized as supporting multiple source use in a number of ways. First, as prior knowledge is foundational to single text comprehension (Kintsch, 1988), it aids students in understanding individual texts in a document set, which in turn supports

multiple text comprehension and corroboration (Le Bigot & Rouet, 2007). McNamara, Kintsch, Songer, and Kintsch (1996) suggest that prior knowledge may be particularly facilitative when texts presented to students are poorly constructed, requiring learners to rely on their prior knowledge to form a coherent understanding of information presented. Prior knowledge may also support students' inference generation (i.e., connection of text-based information with prior knowledge, McNamara et al., 1996) and source integration, when texts are not explicitly linked.

More generally, learners with higher levels of prior knowledge have been found to be more capable of engaging in high-level strategy use (e.g., relevance determinations, source evaluation based on document information) when encountering multiple texts (e.g., Le Bigot & Rouet, 2007; Rouet et al., 1996; Wineburg, 1991). This may be because the cognitive demands posed by processing individual texts in a document set are reduced for high-knowledge students, leaving available cognitive resources for higher level processing, or because high-knowledge students have access to more source use strategies. Finally, high levels of prior knowledge in a particular domain may offer students discipline-specific understanding of document types, content, and structure which may facilitate interactions with diverse texts in a multiple source use context (Gil, Bråten, Vidal-Abarca, & Strømsø, 2010a; Rouet, Favart, Britt, & Perfetti, 1997).

Domain general source evaluation behaviors. The internal resources that students bring to bear on their interactions with multiple texts include both conceptual and procedural knowledge. While conceptual knowledge has typically been assessed as prior knowledge at the topic level, procedural knowledge has been investigated as well. Procedural knowledge of multiple source use has been most commonly considered by

asking students to report the frequency with which they engage in a variety of source use behaviors (Bråten & Strømsø, 2006). For instance, Mason, Boldrin, and Ariasi (2010b) used an *experience in online searching* questionnaire to ask students to rate the frequency with which they use the Internet to research information on a variety of topics (i.e., sports, news).

Considering the role of students' domain-general source use behaviors or strategic practices when using multiple sources is particularly important given that MSU tasks are considered to be more cognitively complex, thereby requiring more strategy use, than single text tasks (Bråten & Strømsø, 2006). More frequent adoption of multiple source use strategies may facilitate using such strategies in less effortful and more effective and flexible ways. Further, the multiple source use strategies learners adopt are considered to be contingent on the task model developed (Strømsø, Bråten, & Samuelstuen, 2003). As such, having greater familiarity or experience with a wide repertoire of strategies may allow students to better meet a broader range of task demands. In the present study, students' reported frequency of engagement in domain general source evaluation behaviors was examined.

Students' experiences employing domain general source evaluation behaviors were the particular focus of this study for a number of reasons. For one, source evaluation has been emphasized as the MSU process most taxed and demanded in online contexts (Coiro, 2003a; Grimes & Boening, 2001; Kammerer & Gerjets, 2012b; Metzger et al., 2003). For another, students' source evaluation behaviors have been thought to be foundational to high-level multiple text processing and, more specifically, multiple text integration (Britt & Aglinskias, 2002; Britt et al., 1999). In reconciling discrepancies

across texts, necessary for multiple source integration, students have been thought to evaluate the validity of specific claims across texts and to determine which accounts to trust (Perfetti et al., 1999). Determining texts' trustworthiness requires students to make judgments of source reliability and authoritativeness and to corroborate information across sources, in other words, to engage in source evaluation. Students with greater prior experience in source evaluation, may be habituated to judging texts and do so more readily across contexts. These students may also be more adept at determining the trustworthiness of a greater variety of sources.

Stance. More recently, attitudes toward task topic have been investigated as impacting students' selections, processing, and evaluations of texts. The role of students' attitudes in single text comprehension has long been recognized in the persuasion literature (e.g., Buehl, Alexander, Murphy, & Sperl, 2001; Murphy & Mason, 2006). Students' attitudes may have particular bearing on multiple source tasks, as multiple texts are considered to be a powerful avenue for presenting controversial information about which there are conflicting points of view (Britt & Aglinskias, 2002; Rouet et al., 1997) and about which students may hold differing beliefs.

Attitudes are considered to impact multiple source use variously. Generally, students holding strong attitudes about a topic may be susceptible to a *confirmation bias* (Nickerson, 1998), selectively attending only to information that is consistent with or supportive of their point of view. Confirmation bias may impact both students' information seeking, wherein students may be predisposed to selecting sources or information they expect to agree with, and their information interpretation and evaluation, wherein students may judge confirmatory evidence more positively

(Nickerson, 1998). Indeed, students have been found to evaluate arguments consistent with their point of view more favorably while being more critical of arguments that conflict with their position (Lord, Ross, & Lepper, 1979; Taber & Lodge, 2006). Having a confirmation bias, in turn, may lead students to engage in one-sided *case-building*, or attending only to information conforming prior beliefs without consideration of conflicting or opposing information (Alexander, Murphy, Buehl, & Sperl, 1998).

Nickerson (1998) further draws the distinction between motivated and unmotivated case-building. When students hold strong attitudes they may be motivated to case build or attend to such evidence and information that supports their position. Alternately, case-building may be unmotivated and occur without students' awareness of the limitations in their reasoning. The potential of motivated case-building to emerge suggests the need to consider not only students' specific views on an issue but also the valence or strength of students' attitudes. In fact, strength of attitudes, or attitude certainty, has been found to be more predictive of behavior than specific attitudes in-and-of themselves (Nan, 2009; Taber, Cann, & Kuscova, 2009). In addition to prompting case-building, strong attitudes may serve a motivational purpose, increasing task engagement.

Beyond affecting the process of students' source use, attitudes may impact product as well. When asking students to advise a fictitious friend about whether or not to take cholesterol medication, Kienhues, Stadtler, and Bromme (2011) found that after reading information consistent with their prior position, either in support of or against prescribed medication, participants retained their previous position and felt more certain in their decisions. In the present study, rather than examining attitudes *per se* students'

stance with regard to the topic of the multiple source use prompt, the Arab Spring in Egypt, was considered.

Specifically, students were asked to position themselves with regard to the target prompt, asking who should hold power in Egypt (i.e., Mohamed Morsi, General el-Sisi, or an “other” option), and to report their degree of commitment to such a stance.

Attitudes, *per se*, were not fully examined as these have been considered to correspond to well-developed and coherent systems of beliefs accompanied by “extensive, well-organized knowledge structures” (Wood et al., 1995, p. 284). In the present study, based on pilot data collected and the sample’s limited prior knowledge, participants were not expected to hold strong attitudes with regard to the target prompt. However, as the topic selected represented a controversial, contemporary political issue, students’ stances, or initial positions with regard to the task, were examined. McCrudden and Sparks (2014), similarly, in examining students’ beliefs with regard to a contemporary topic asked learners to position themselves in reference to the assigned task (i.e., in support of or in opposition to widening the Victoria tunnel in New Zealand).

Despite the documented effects of individual difference factors on students’ source use, these have been examined to a limited extent in the multiple source use literature. For instance, prior knowledge has most commonly been used as a control factor in studies of multiple source use (e.g. Strømsø, Bråten, & Britt, 2010) rather than explored as impacting the MSU process and product. Thus, a secondary purpose of this study was to consider the nature of the relations between students’ cognitive resources (i.e., prior knowledge, stances, domain general source evaluation behaviors) and source use behaviors (i.e., source selections, source processing or time on texts, source

evaluations), within the framework introduced by the MD-TRACE model. Finally, this study examined the extent to which individual difference factors in conjunction with source use behaviors impacted response quality when students completed a multiple source task.

Prior knowledge, students' stance on the target issue, and domain-general source evaluation behaviors were the individual difference factors selected for examination for a number of reasons. First, prior knowledge has been the individual difference factor identified as having the greatest role in text processing and extensively examined in both single and multiple text use (Alexander et al., 1994; Alexander et al., 1991; Kerstetter & Cho, 2004). It is also the internal cognitive resource predominantly specified by Rouet and Britt (2011) as bearing on the processes identified in the MD-TRACE model. Further, as the MD-TRACE model is considered to be behavioral in nature, students' prior experience with multiple source use, a behavioral measure, was considered to be important to include. To the extent that students' source evaluation has been the source use process most extensively examined in prior research due to its foundational role in multiple text integration (Britt et al., 1999), students' experience with or frequency of engagement in domain general source evaluation behaviors was the behavior-based learner characteristic of primary interest in this study.

Students' stance with regard to the target prompt was included because the topic of the multiple source task addressed a contemporary political issue that was specifically selected to be *controversial*, with strong and discrepant views presented by each side and students asked to align with (Britt & Aglinskis, 2002). Given its controversial nature, students' stance, or initial position and strength of affect with regard to the target prompt

was considered to be an important factor to examine.

The conceptual model, including both learner characteristics and multiple source use behaviors, examined in this study is presented in Figure 2. Like the MD-TRACE model, the conceptual model introduced focuses on students' source processing behaviors (i.e., source selection, source processing, and source evaluation) and their effects on response quality. Further, three learner characteristics are explicitly specified (i.e. prior knowledge, stance, and domain general source evaluation behaviors), and their effects on multiple source use behaviors and on ultimate response quality are modeled.

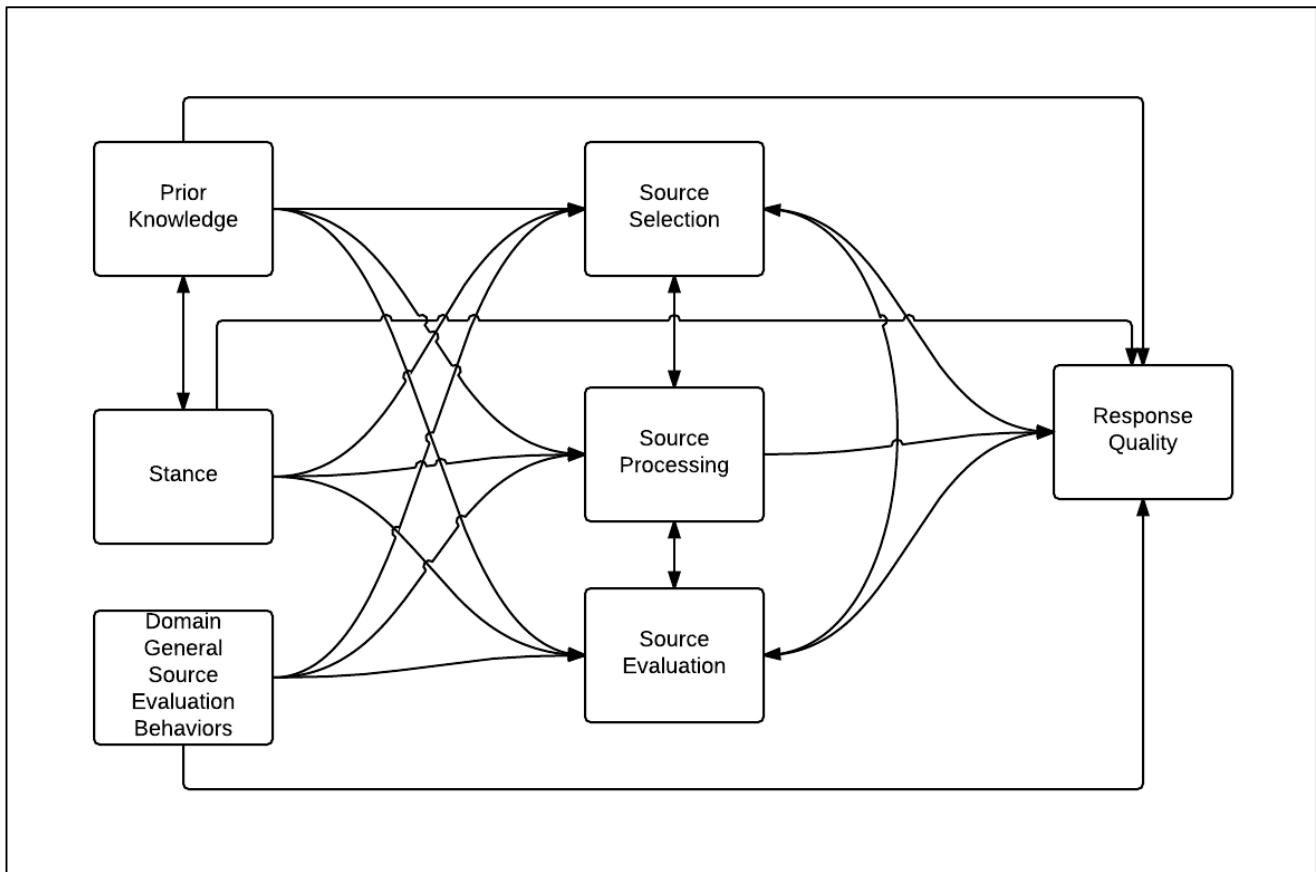


Figure 2. Components of the conceptual model of individual difference factors and behaviors in multiple source use.

Gaps in the Literature

Two specific gaps in the literature were addressed in the present investigation. First, a comprehensive model of students' multiple source use behaviors was examined to determine the extent to which students' engagement in source selection, processing, and evaluation influenced response quality. Such analysis is in response to Rouet and Britt's (2011) call for a specific and comprehensive examination of their proposed model. Second, individual difference factors (i.e., prior knowledge, students' stance on the issue, and domain general source evaluation behaviors) were considered as they were associated with source use behaviors (i.e., source selection, processing, and source evaluation) and predictive of response quality. The effects of a joint model, including both individual difference factors and multiple source use behaviors, on task performance was the central focus of this study. Such a model, encompassing both core source use behaviors and individuals' cognitive resource, has yet to be fully investigated.

The Current Investigation

To address the identified gaps in the literature, data were collected in two sessions. In Session 1, students were asked to complete a prior knowledge measure and report their stance about the chosen topic (i.e., Arab Spring in Egypt) and the frequency with which they engaged in a variety of behaviors associated with source evaluation. In Session 2, students were asked to engage in a multiple source use task, or research a prompt using a library of sources, and to compose a written response. While students researched information in the source library, trace data of their source use behaviors were collected. Specifically, data were gathered on source selections (i.e., number of texts students access), source processing (i.e., time on texts), and source evaluations (i.e., whether or not document information was accessed about each source used; source

ratings along a variety of dimensions).

The data collected mapped on to factors introduced in the conceptual model presented in Figure 2. Specifically, in Session 1, students were asked to report their prior knowledge, domain general source evaluation behaviors and stance on the target issue. These were the individual difference factors identified as contributing to task performance in the conceptual model guiding this work.

In terms of multiple source use behavioral factors, data were collected and mapped onto each multiple source use behavior of interest (i.e., source selection, source processing, and source evaluation). The total number of texts students accessed was the indicator of source selection. The total time students devoted to source use was considered to be a measure of text processing. Finally, four indicators were used to capture students' source evaluation. These considered (a) whether or not students accessed document information about each source selected and (b) students' ratings of sources used along key dimensions (i.e., trustworthiness, usefulness, and interestingness). This set of source evaluation metrics was selected to reflect the ways in which students' text judgments have been examined in prior research (e.g., Bråten et al., 2009; Bråten et al., 2011; Britt & Aglinskias, 2002; Rouet et al., 1996).

Students' ratings of source *trustworthiness* have been most commonly investigated in the literature (e.g., Bråten et al., 2009; Bråten et al., 2011). Trustworthiness evaluations are considered to underlie high-level text use, as multiple text integration is considered to be a process of corroborating competing information across sources and determining which to trust (Britt, Perfetti, Sandak, & Rouet, 1999). While some researchers have asked learners to make general judgments associated with

source trustworthiness (Stadtler & Bromme, 2008), Perfetti et al. (1999) consider trustworthiness evaluations to be based on students' considerations of each source's document information (e.g., author, credentials, publisher). At the same time, students have been found to be limited in accessing and using document information (Britt & Aglinskias, 2002; Wineburg, 1991) or engaging in *sourcing*. As a result, students' competencies and frequency of engagement in sourcing have received considerable attention in the literature (Britt, Weimer-Hastings, Larson, & Perfetti, 2004; Perfetti, Rouet, & Britt, 1999; Stahl et al., 1999). In the present study, whether or not learners elected to *access document information* for each source they selected was examined, as were students' ratings of source trustworthiness rendered after using each source.

While trustworthiness evaluations have received most attention in the literature, Rouet and Britt (2011) suggest that students' evaluations or general perceptions of source *usefulness* are key, preceding ratings of source trustworthiness. Ratings of source usefulness are considered to be evaluations of a text's pertinence or instrumental value in meeting task goals. As multiple source use is considered to be a task-driven process (McCrudden & Schraw, 2007), usefulness evaluations may be the primary judgments students render when accessing a text. Indeed, Rouet and Britt (2011) suggest that students likely first determine whether a source is useful in meeting task goals, prior to engaging in the more cognitively demanding process of deciding source trustworthiness. In this study, as in prior work, students were asked to evaluate sources along both usefulness and trustworthiness dimensions (Britt et al., 1999; Rouet et al., 1996; Rouet et al., 1997)

In the present analyses, an additional source evaluation dimension was

introduced. Students were asked to evaluate the *interestingness* of each source accessed. Interest has been considered to be a key motivational variable in both single (Alexander, Kilikowich, & Jetton, 1994, for a review see Hidi & Baird, 1986) and multiple text processing (Strømsø, Bråten, & Britt, 2011). Text interestingness has been found to correspond to text recall, comprehension, and elaboration, resulting in deeper text processing (Hidi, 2001; Hidi & Baird, 1988; Krapp, 1999; Schiefele, 1999; Strømsø et al., 2011). It has also been found to be a key motivational variable supporting effortful engagement and persistence in task completion (Hidi, 1990; Mason & Boscolo, 2004), particularly when challenges are experienced. Given the cognitive complexity posed by multiple source tasks (Bråten & Strømsø, 2006), text interestingness was considered to be an important evaluative dimension to include as potentially predictive of students' task performance. While in prior research, interest has most commonly been assessed at the topic level, prior to task engagement (Bråten, Gil, Strømsø, & Vidal-Abarca, 2010; Strømsø, Bråten, & Britt, 2010), in this study, learners were asked to rate the interestingness of each text they accessed. Source interestingness or students' ratings of text-based interest (Hidi, 1990) was considered to be a preferable measure of situation-specific motivation than was learners' topic interest. Evaluations of text interestingness captured students' engagement throughout the task, while processing each source in the library.

Based on the conceptual model displayed in Figure 2, the model presented in Figure 3 was investigated.

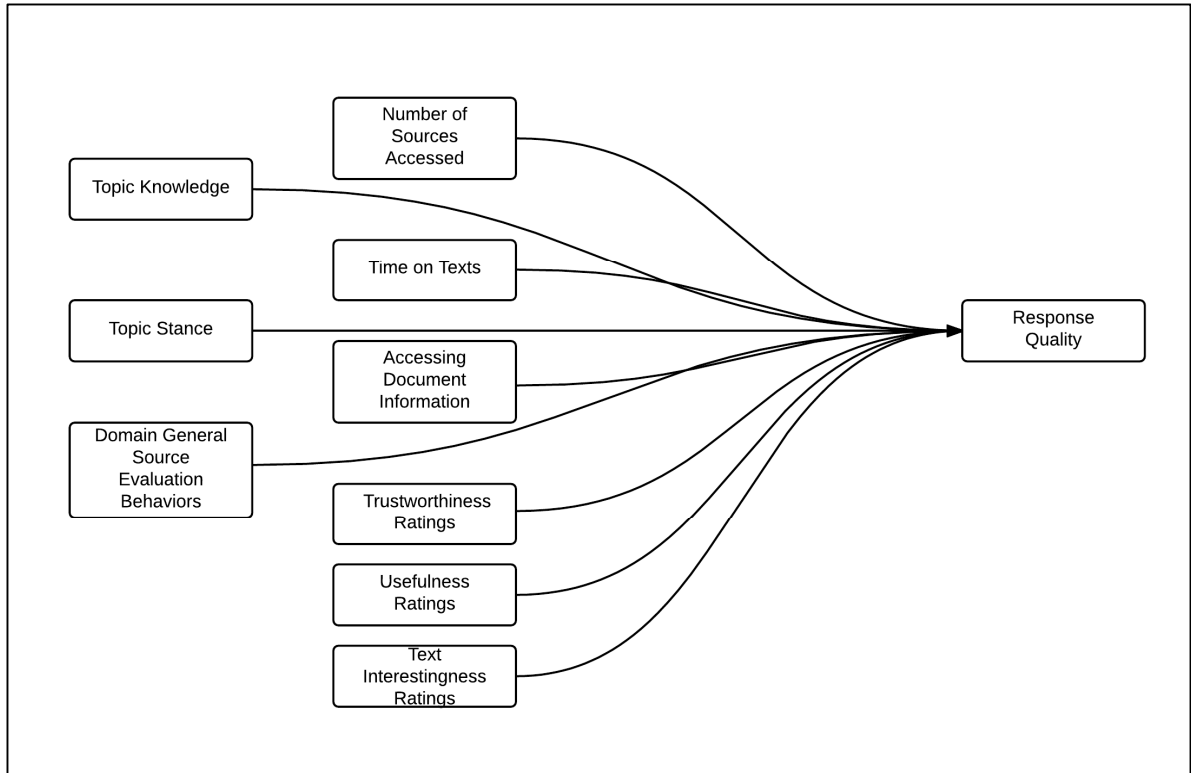


Figure 3. Model of learner characteristics and multiple source use behaviors predicting open-ended response quality, examined in this study.

Research Questions

Five research questions were addressed.

1. What is the nature of students' multiple source use when responding to a multiple source use task?
2. To what extent do source use behaviors (i.e., number of sources selected, time on texts, and source ratings) predict response quality when students complete a multiple source task?
3. What is the nature of the association between individual difference factors (i.e., prior knowledge, topic stance, and domain general source evaluation behaviors) and students' manifest source use behaviors (i.e., number of sources selected, time on texts, and source ratings)?

4. To what extent do individual difference factors (i.e., prior knowledge, topic stance, and domain general source evaluation behaviors) predict response quality when students complete a multiple source task?
5. To what extent do individual difference factors (i.e., prior knowledge, topic stance, domain general source evaluation behaviors) and multiple source use behaviors (i.e., number of sources selected, time on source use, and source ratings) predict response quality when students complete a multiple source task?

Key Terms

The following terms were central to the conduct of this investigation:

Attitudes: Attitudes may be defined as, “general evaluations individuals have regarding people, places, objects and issues” (Petty & Brinol, 2010; p. 217). Attitudes are considered to be composed of a positive or negative evaluation of an object or issue (i.e., value) as well as the strength of or commitment to this evaluation (i.e., valence, Petty & Wegener, 1997).

Credibility: Credibility is the determination of source believability, made based on author expertise or authoritativeness and source trustworthiness (Hovland, Janis, & Kelley, 1953).

Document Information: Document information is *metadata* or a sources’ identifying information related to its purpose for being created or origin. This includes information about the author(s), intended audience, and when/where a text was published (Goldman & Scardamalia, 2013a).

Document: A document is a term that emphasizes texts as *social artifacts*. The term document reflects a conception of texts that recognizes authors' goals, desired audience, and purpose for writing as integral to understanding content. Document refers to a broad conception of texts, including its content as well as information about author, context and setting, and document information (Britt, Rouet, & Braasch, 2013).

Documents are defined in disciplinary terms as meeting a particular domain's standards and conventions for writing (Goldman & Scardamalia, 2013a).

Document/Source Type: Source type refers to text genres or kinds of sources that follow disciplinary conventions. Source types are "forms of documents that have identifiable elements, rules of form, and content" (Dillon & Gushrowski, 2000, p. 202), such as textbooks, newspapers, or research studies.

Epistemic Beliefs: Epistemic beliefs are beliefs about knowledge and knowing (Hofer & Pintrich, 1997). Dimensional conceptions of epistemic beliefs have identified four dimensions of such beliefs: students' beliefs about *sources of knowledge*, *justifications for knowledge*, *certainty of knowledge*, and *simplicity of knowledge* (Strømsø, Bråten, & Britt, 2011).

Epistemic Dimensions of Source Evaluation: Epistemic dimensions of source evaluation are judgments of sources considered to be manifestations of students' underlying epistemic beliefs (e.g., judgments of source authority, scientific nature, Mason, Ariasi, & Boldrin, 2011; Mason et al., 2010a, 2010b). These are source evaluations aimed at establishing knowledge or truth (e.g., judgments of source trustworthiness, Bråten et al., 2009; accuracy, Brem, Russell, & Weems, 2001).

Evaluation: Evaluations are judgments of or conclusions about sources or the information within them based on “available or accessible information about the source” (Bråten et al., 2009, p. 6). Evaluation statements “convey positive or negative judgments about some aspect of the text” (Wolfe & Goldman, 2005, p. 480).

Information Source (Source): A source is a generic form of document or *text writ large* (Goldman & Scardamalia, 2013a). The origin of information in a source may be print, digital, visual, or auditory.

Interest: Hidi (2006) defines interest as a “unique motivational variable, as well as a psychological state that occurs during interactions between persons and their objects of interest” (p. 69). Interest has been considered to take two forms: *individual interest*, representing a sustained personal predisposition toward content, and *situational interest*, referring to momentary interest arising from contextual factors.

Interestingness: Interestingness or text-based interest refers to situational interest that is generated through students’ interactions with texts (Garner, Alexander, & Gillingham, 1991; Hidi, 2006).

Justifications: Justifications refer to criteria or reasons by which students form broad source evaluations. For instance, justification criteria for an evaluation of source trustworthiness have included reasons related to author, publisher, and source type (Bråten et al., 2009).

Multiple Source Use (MSU): MSU refers to students’ engagement “in the processes of search, selection, evaluation, comparison, and integration of ideas from multiple sources of information” to complete academic tasks or solve problems (Wiley et al., 2009, p. 1061).

Non-Epistemic Dimensions of Source Evaluation: Non-epistemic dimensions of source evaluation are judgments of sources not associated with students' epistemic beliefs and not concerned with knowledge establishment. Judgments of task-related (i.e., relevance) factors or surface source features (e.g., formatting, Rouet, Ros, Goumi, Macedo-Rouet, & Dinet, 2011) are considered to reflect non-epistemic dimensions of source evaluation.

Prior Knowledge: Prior knowledge is composed of domain and topic knowledge. Alexander, Schallert, and Hare (1991) characterize domain knowledge as the breadth or span of subjects' knowledge about a field, whereas topic knowledge may be characterized as the depth of a subjects' knowledge about a particular topic.

Relevance: Judgments of relevance are students' determinations that a source serves an instrumental purpose in meeting their goals for source use or judgments of *task-based importance* (Rouet & Britt, 2011). Relevance judgments may be either positive (i.e., related to task) or negative (i.e., unrelated to task), and may pertain to texts as a whole (i.e., source relevance) or to specific content (i.e., information relevance, Anmarkrud, McCrudden, Bråten, & Strømsø, 2013).

Source Evaluation Behaviors: Source evaluation behaviors refer to actions students may engage in to ascertain source credibility (Metzger et al., 2003). Specifically, these are behaviors associated with establishing the “accuracy, authority, objectivity, currency, and coverage” of a source (p. 2079).

Source Selection: Source selection refers to accessing a text from a collection of sources presented as a set (e.g., via a hyperlinked library or search engine results page).

Often selection occurs based on only limited document information presented about each text (Gerjets, Kammerer, & Werner, 2011).

Sourcing: Sourcing, a source evaluation heuristic identified by Wineburg (1991), is defined as, “looking first at the source of the document before reading the text itself to consider how the bias of the source might have affected the content of the document” (Stahl, Hynd, Britton, McNish, & Bosquet, 1996, p. 433). In effect, sourcing is the process of attending to any document information that may aid in source evaluation (Britt & Aglinskas, 2002).

Stance: Stance refers to students’ positioning or adoption of a view in reference to an issue or topic. While attitudes have typically been associated with well-formed and deeply held beliefs (Wood, Rhodes, & Biek, 1995), stance represents a more shallow commitment to a position.

Text: Text is defined as connected, written information that is coherent and continuous. Text consists of a microstructure, at the level of individual statements in a sentence, and a macro-structure, referring to the overall connections in a text (Kintsch & van Dijk, 1978).

Text Processing: Kintsch and van Dijk (1978) define text processing as including three sets of operations that readers may perform on texts, simultaneously or sequentially. These operations include (a) organizing textual elements into a coherent whole, (b) reducing information in texts to a *gist*, and (c) creating new texts, which are texts, reconstructed based on readers’ prior knowledge.

CHAPTER 2: REVIEW OF THE LITERATURE

In this review of the literature, the Multiple Documents Task-Based Relevance Assessment and Content Extraction (MD-TRACE) model of multiple source use is introduced as the framework guiding the present investigation. Then, the core source use behaviors specified in the MD-TRACE model (i.e., source selection, source processing, and source evaluation) are considered and their assessment discussed. Next, individual difference factors (i.e., prior knowledge, domain general source evaluation behaviors, and topic stance) pertinent to multiple source use are examined. Construct definitions of each learner characteristic and their potential impact on multiple source use are explained. Finally, this chapter concludes with a review of findings from initial examinations of models of multiple source use.

Given that multiple source use has been conceptualized as developing throughout students' schooling (Britt & Aglinskias, 2002), this literature review focuses primarily on investigations of upper secondary and college level samples, as these are most informative in framing the present study, targeting undergraduate students.

Model of Multiple Source Use

Multiple source use has been defined as students' engagement "in the processes of search, selection, evaluation, comparison, and integration of ideas from multiple sources of information" to complete academic tasks or solve problems (Wiley et al., 2009, p. 1061). These processes have been identified as critical for the development of literacy in the 21st century and for post-secondary success (Common Core State Standards Initiative, 2010; Goldman & Scardamalia, 2013b). Yet limited models exist capturing the dynamic processes involved in multiple source use.

Britt and Rouet (2011) describe the need for a model of multiple text comprehension for cases when students are asked to solve open-ended problems or to learn about complex topics. In such situations, single texts may prove to be insufficient; rather, students may be required to consult multiple texts. In fact, models of multiple source use are thought to examine students' processing of *complex documents*, or documents that "include more than one piece of coherent, continuous text" (Rouet, 2006, p. xvii). This in turn means that models of multiple source use investigate the cognitive processes associated with interacting with such documents. In models of multiple source use, the focus shifts from understanding students' more limited comprehension of individual texts, to capturing students' abstract representations of complex situations and phenomena described across sources (Rouet, 2006). Britt, Perfetti, Sandak, and Rouet refer to this as "increasing the grain size" of comprehension to be examined (1999, p. 210).

The MD-TRACE model has been proposed as a "generalized theory of multiple text processing" (Rouet & Britt, 2011, p. 3). This model builds on earlier conceptions of information search that consider how students locate specific pieces of information across texts (e.g., Guthrie, 1988) to identify the cognitive processes implicated when students use information in multiple texts to form comprehensive understandings of complex issues (Rouet, 2006). The MD-TRACE model defines the internal cognitive and external resources needed to support multiple source use and outlines the core processes involved: construction of a task model, determination of an information need, selection, processing, and integration of information across documents, creation of a task product, and determination of the correspondence between task product and task demands. Three of

these processes (i.e., source selection, processing, and integration) are the particular focus of this investigation as is the role of students' internal resources bearing on multiple source use.

The MD-TRACE model is considered to be one of *functional reading*, wherein students engage with texts in order to meet specific goals (Rouet & Britt, 2011). Rouet and Britt identify five phases or steps students engage in to satisfy such goals. These five steps are delineated in the following overview.

Step 1: Task Model Construction

In the first step, students develop a *task model*, or cognitive representation of task demands. Three task-related components are included in a task model: students' understandings of task goals, an outline of procedures that may be undertaken to achieve these goals, and criteria along which satisfied task goals will be evaluated (Rouet & Britt, 2011). As such, the task model determines the goals for multiple source use and gives shape not only to the process of multiple source use but to the outcome as well. The formation of a task model is based on readers' prior knowledge and experiences with similar task (i.e., internal resources) and on the affordances and constraints provided by the task context (i.e., external resources, Rouet & Britt, 2011).

Step 2: Determination of Information Need

In Step 2 of the MD-TRACE model, students make a determination of *information need*, or decide, based on their task model, that they have insufficient information to meet task demands and thus elect to access multiple information sources. Determining an information need maybe a *general* decision, when students decide that they have insufficient information about a particular topic, or *specific*, when students look

for particular information in texts. Information need determination is considered to be an iterative process. As students gain information about a particular topic they make successive judgments about how much and what types of further information may be needed (Rouet & Britt, 2011). As such, determinations of information need are contingent on two types of judgments: students' assessments that they need information to meet task demands and decisions that they have sufficient information to compose a response (i.e., that information needs have been met).

Step 3: Source Processing

Step 3 of the MD-TRACE model, source processing, is composed of three sub-steps: source *selection*, *processing*, and *integration*, dependent on source *evaluation*. These processes, while independent, are considered to be sub-steps of Step 3 because they are highly interrelated and dependent on one another. These three processes are considered to be most complex, in part, because they occur both within individual texts and across documents. Source selection, processing, and integration are considered to be guided by continuous *relevance assessments*. These are judgments of sources' appropriateness for meeting task demands, as well as for learners' needs and capabilities (McCrudden & Schraw, 2007; Rouet & Britt, 2011).

Relevance assessments may be construed broadly, extending to include students' considerations of *source trustworthiness* and *usefulness* (Rouet & Britt, 2011). For instance, if a task model calls for information to be garnered from reliable sources, students may judge unreliable texts to be irrelevant in meeting task demands. As such, trustworthiness ratings allow students to ascertain whether a given text constitutes a sufficiently reliable or authoritative source of information for responding to task demands

(Mason et al., 2010b). Rouet and Britt (2011) suggest that students evaluate source trustworthiness based on document information (i.e., identifying source information like author or publisher) and through corroboration (i.e., by comparing information consistency across texts). Source usefulness is determined based to task demands and reflects the pragmatic value of a source in satisfying task goals. While students' relevance judgments, including trustworthiness and usefulness evaluations, have commonly been examined when students are selecting texts (e.g., Gerjets et al., 2011; Rouet et al., 2011), the MD-TRACE model suggests that source selection, processing, evaluation, and integration are all based on implicit judgment of relevance.

Source selection. More than other processes in the MD-TRACE model, Rouet and Britt (2011) suggest that source selection occurs primarily based on relevance determinations. In selecting sources, students are thought to make decisions regarding a source's presumed *topic* relevance and *task* relevance. Topic relevance refers to the content overlap or lexical and semantic correspondence between texts and the task. Task relevance is considered to be a decision about the alignment of a text's or author's goals with students' task goals (Rouet & Britt, 2011).

Relevance judgments are also considered to be students' cost-benefit assessments of how much a document may contribute to helping them meet task goals vis-a-vis how much effort it may take to access and process the document (Britt & Rouet, 2011; Wilson & Sperber, 2002). Such a definition of source relevance highlights the cost-benefit nature of students' decisions about source selection. When accessing sources, learners make decisions regarding whether particular documents are "worth" the effort needed to process them. As such, learners' choices with regard to source selection are comparative

in nature. Rather than selecting sources in the absolute, students make decisions about source access in reference to the perceived contributions and costs of use associated with other texts in a document set.

Source processing. Once a source is selected, students' engage in text processing or single text comprehension, before integrating information found in a particular source with other texts. Models of multiple text comprehension have been based on Kintsch's (1998) Construction Integration Model for single text comprehension. Kintsch suggests that understanding a single text occurs through students' construction of two cognitive models, the *text-based* and *situation* models. While the text-based model is a propositional model, representing the local and global meaning of texts, the situation model, integrates information in a text with students' prior knowledge to develop comprehensive and connected understanding. While the text-based model may be considered to be text-internal, the situation model is text-external, connecting with students' prior schema and information beyond the text base (Strømsø, Bråten, & Samuelstuen, 2003). The Construction Integration model, therefore, specifies a process whereby students *construct* a text-based model and *integrate* information from text into their prior knowledge to build situation model understanding.

In constructing these cognitive models, learners are thought to alternate between the complementary and iterative processes of surface-level scanning and deep level processing (Rouet & Britt, 2011). While the former has been considered in studies of information location (e.g., Payne & Reader, 2006), the latter has been examined in models of single text reading and comprehension. For example, Beizhuizen and Stoutjesdijk (1999) distinguish surface and deep level processing as differentiated by

students' goals during source use. Specifically, when engaged in surface level processing students are driven by a desire to recall as much information as possible, thereby engaging in linear text processing, thorough reading with attention to detail, and rehearsal and memorization. Deep processing is marked by efforts not only to remember a specific text but also to use text to understand a particular topic. When processing deeply, learners approach texts more globally, relate aspects of text to one another, organize information and relate it to their prior knowledge, and adopt a critical approach to text processing. As such, surface processing may be considered to be *serialistic* and deep processing to be *holistic* (Pask, 1976). Beizhuizen and Stoutjesdijk (1999) emphasize the need for both deep and surface level strategies and flexibility in their adoption. They also discuss the difficulty in drawing a firm distinction between these, suggesting that summative indicators may be appropriate to capture students' single text processing.

In part, the depth of processing (i.e., surface level scanning or deep level processing) students engage in is determined by task. For instance, locating specific pieces of information in text may require students to only scan sources, identifying and extracting needed facts (Guthrie & Kirsch, 1987). More open-ended or complex tasks may demand deeper processing to comprehensively understand issues described in texts (McCrudden, Schraw, & Hartley, 2006; Rouet & Britt, 2011).

Source evaluation. The final sub-process identified in Step 3 of the MD-TRACE model, source integration, has been modeled in its own right and is considered to be contingent on students' source evaluations (Rouet & Britt, 2011). Multiple source use requires that once students have acquired information from a particular source, that information is somehow combined or integrated with information from across additional texts. This integration process is complicated by the variety in the types of relations between and among texts that can be identified. Sources may add to each other, support each other, oppose, or contradict one another (Rouet & Britt, 2011). As such, multiple source use is not just a process of aggregating information but rather requires the corroboration, reconciliation, and ultimate integration of information presented across sources.

Multiple text integration is dependent on students' source evaluations (Perfetti, Rouet, & Britt, 1999; Rouet, 2006). For instance, when two texts conflict, students must decide which to believe. They might do this in part by evaluating source trustworthiness and author expertise. The multiple source integration process and its dependence on source evaluation are articulated in the more specific Documents Model of Multiple Source Use. The Documents Model, further expanded upon in the subsequent section, presents multiple source integration as reliant on source evaluation. In the present study students' source evaluations are examined as they support subsequent integration, assessed via open-ended response quality.

Step 4: Response Formulation

In Step 4 of the MD-TRACE model students develop a textual product in response to task demands. This process may be considered to be a transformation,

whereby students reformulate information presented across documents into a desired product. The nature of the task determines the extensiveness of the transformation demanded. For instance, locating and reproducing a specific piece of information in text requires little transformation as compared to generating a more elaborated response (Mosenthal, 1996). The textual products students develop are based on their task models or understandings of task expectations and constraints.

Evidence suggests that students find response composition to be quite challenging. For instance, assessments of graduating secondary students have revealed that only a minority of students are able to craft arguments based on information presented in texts (NAEP, 1996, 1998). Students may find producing responses based on conflicting or opposing texts to be particularly challenging (Wolfe & Britt, 2008; Wolfe, Britt, & Butler, 2009; Wolfe & Goldman, 2005). Just such conflicting content is common in multiple source use tasks (Britt & Aglinskias, 2002). Problems have likewise been identified in students composing one-sided essays, without considering alternate positions, or in failing to elaborate arguments (Rouet & Britt, 2011).

Step 5: Response Evaluation

The fifth and final step of the MD-TRACE model involves determining whether a text product meets task demands. In other words, students evaluate their responses produced in Step 4 according to the criteria or task expectations set out in Step 1, with their construction of a task model.

Internal and External Resources

The MD-TRACE model extends beyond considering students' behaviors when engaged in multiple source use to also examine learner and context-specific factors that

may affect multiple source use. These learner and context-related factors have been termed internal and external resources, respectively, and are thought to have an impact on students' interactions with texts. Internal or cognitive resources, such as prior knowledge, are factors that are nested in students and that learners are thought to bring with them to the task context. Internal resources may be either permanent or transitory. Although prior knowledge may be considered to be a fairly stable resource contributing to students' multiple source use (i.e., permanent), the task model students construct is a transitory resource that is specific to the particular demands of a given task.

Internal resources are considered to be composed of both declarative and procedural knowledge. They include both the prior topic knowledge (i.e., declarative) and text use skills (i.e., procedural) that differentiate students' task performance. Both declarative and procedural knowledge are considered to have a bearing on multiple source use processes. For instance, higher levels of prior knowledge may result in students being able to comprehend more sources, thereby enabling them to more effectively process more texts within a document set. Greater task experience may aid students in constructing a task model or more effectively composing a textual product (Rouet & Britt, 2011). The internal resources to be examined in the present study include students' prior knowledge and stance toward the topic (i.e., declarative knowledge), as well as students' reported domain-general source evaluation behaviors (i.e., procedural knowledge).

External resources, such as texts included in the source library, are affordances particular to the task context. Three types of external resources have been identified. These may be presented either prior to text engagement or during source use. The first

are task specifications, including the task or prompt itself as well as any associated directions. Task specifications may also be indirect or non-explicit, such as time constraints or the particular conditions under which the task is presented. The second type of external resource is the information or content included as part of the task per-se. Rouet and Britt (2011) identify not only the information presented in texts as belonging to this category but also information presented *about* texts (i.e., document information such as author or publisher) as constituting this second type of external resource. The third type of external resources is the learner-generated task product, or students' responses to a given task. This third type of external product is thought to be an integration of both task specifications and text information.

While the MD-TRACE model sets out the processes involved in multiple source use, a complementary and nested model, the Documents Model of Multiple Source Use (Britt et al., 1999), has been introduced to specifically explain how students complete one of these processes, source integration. The Documents Model considers multiple source integration to be dependent on source evaluation. While the MD-TRACE model is primarily focused on students' relevance evaluations (i.e., students' judgments of information as consistent with their needs and capacities, McCrudden & Shraw, 2007), the Documents Model is focused on students' evaluations of source trustworthiness.

Documents Model of Multiple Source Use

Based on the Construction Integration framework introduced by Kintsch and van Dijk (1978), the Documents Model of Multiple Source Use (Britt et al., 1999) describes students' integration of information across multiple documents, as indicated in Step 3c of Rouet and Britt's (2011) MD-TRACE model. The documents model specifies the comprehension of multiple texts as occurring through students' construction of two cognitive models, the *situation* and the *inter-text* models; models that represent the relations of information within and across texts.

According to the documents model, when drawing on multiple sources of information, students must build connections between them to develop an integrated understanding of the common situation described among them. The *situation model* is students' broad representation of the common situation or information described in the texts, while the more complex *inter-text model* represents the relation between each document and elements of the situation as well as the relations among the various documents (Perfetti et al., 1999). The situation model is comprehensive and includes information about each individual source, including its content, document information, setting and form, as well as the author's goals and intentions (Britt et al., 1999). The inter-text model is selective and includes information deemed to be most important and relevant. The specific pieces of information included in the inter-text model, referred to as *core events*, are both linked to the specific documents they originate from and integrated across sources. These core events may be repeated across documents or unique to a particular text (Britt et al., 1999).

The documents model has also been referred to as one of *content integration* and *source separation* (Britt et al., 1999), in reference to the dual characteristics of this model that make it the most adaptive approach to comprehending multiple documents for students. The documents model requires the content or situation described across multiple texts to be integrated, while each of the contributing sources is simultaneously kept separate, such that each piece of information included in the integrated model is *tagged* or associated with a particular source. Such a model can be contrasted with less sophisticated approaches to multiple text integration (i.e., other inter-text models), including the *separate representation model*, wherein information is linked to its source of origin but not connected across texts, and the *mush model*, where information is integrated but not connected to the source it originates from.

These less sophisticated models are thought to interfere with either understandings of content and information integration (i.e., separate representation model) or with understandings of sources and the resolution of inconsistencies between texts (i.e., mush model) stemming from a lack of attention to document information (Britt et al., 1999). In particular, the mush model, considered to be the one most commonly produced in school setting where students are primarily motivated to learn about a topic, has been thought to hinder students' engagement in information verification and source evaluation, processes that require sourcing and attention to document information.

For students, the documents model is also considered to be preferable to a fourth possible model of comprehension, the *tag all model*. In the tag all model, each piece of information is linked to its document of origin and connections are made between all information across sources and between all of the sources themselves. The tag all model

offers an exhaustive approach to comprehension and is considered to be a scholar model, too cognitively demanding for non-expert learners (Britt et al., 1999). Unlike with the tag all model, when constructing the documents model, good, non-expert readers integrate information across sources by making decisions regarding which information to incorporate into and which to exclude from their inter-text model (i.e., not all information is included). Although the documents model is more accessible for students than the tag all model, it nonetheless requires learners to make judgments of which information to privilege and which to exclude; these judgments rest on students' attendance to document information and source evaluations.

Sourcing, Source Evaluation, and the Documents Model

The documents model emphasizes the evaluation of sources as necessary for multiple source integration and introduces a theoretical conceptualization of students' source evaluations. When encountering multiple texts, good readers attend to both content and document information (i.e., connect information to its source of origin). Such attendance to document information is necessary not merely to support text comprehension. It is also at the core of source evaluation; in focusing on document information, students draw conclusions about source quality or credibility. Source evaluations become particularly pertinent in developing an integrated inter-text model. When students make decisions about which information to retain, and which to discard, such determinations are based on students' judgments of source quality and reliability (Bråten et al., 2009). Information from sources high in credibility or information that can be verified across texts may be given priority, judged to be more accurate or important, and thus more likely to be included in the inter-text model. Further, when sources

conflict, information is retained from those texts considered to be more reliable (Bråten et al., 2009). Such determinations of source reliability or trustworthiness are based on students' engagement in sourcing (i.e., attendance to or use of identifying document information, such as author or publisher). Considerations of document information are thought to be foundational to source evaluation, in turn impacting source integration.

The documents model puts forth the theoretical notion that students' attendance to source information leads to their source evaluations, or overall text judgments; these judgments then inform students' integration of information across multiple texts (i.e., decisions about which information to include or exclude from their inter-text model). This understanding of source evaluation has been operationalized in a number of studies (Bråten et al., 2009; Bråten et al., 2011). In these empirical investigations, students have been asked to rate the trustworthiness of sources, or evaluate sources overall, and then asked to report the extent to which their source evaluations were based on judgments of various document information, including author, publisher, source type, and date of publication. This methodological approach has been widely adopted in the literature, with students evaluating or rating texts along a variety of dimensions including trustworthiness, usefulness (Rouet et al., 1996), reliability (Wiley et al., 2009), credibility (Hargittai, Fullerton, Menchen-Trevino, & Thomas, 2010) and the justification of evaluations to varying extents.

In specifying students' interactions the multiple sources, in Step 3, the MD-TRACE model would be more complete in including students' attendance to document information and source evaluations as the key sub-processes necessary for both multiple text integration and ultimate response composition. In the present study, the sub-

processes set forth in Step 3 of the MD-TRACE model (i.e., source selection and source processing), as well as sourcing and source evaluation, as necessary for multiple text integration and comprehension, are examined. The next section reviews each of these sub-processes (i.e., source selection, source processing, and sourcing and source evaluation) and their assessment. Then, the relation between each of these identified processes and particular internal resources is explored.

Processes of Multiple Source Use

Three specific multiple source use processes set out in the MD-TRACE model are examined. These are students' source selections, source processing, and source evaluations.

Source Selection

Source selection, in the multiple source use literature, refers to students' accessing of sources from a document set. Documents are typically presented to students through *electronic documentary environments* (Le Bigot & Rouet, 2007). These are digital systems for text presentation that introduce sources alongside one another and include document information (e.g., author) about each text. They are commonly formatted either as a source library (Britt & Aglinskas, 2002), a hyperlinked table-of-contents-like listing of texts (Le Bigot & Rouet, 2007), or as a search engine results page (SERP, Kammerer, Wollny, Gerjets, & Scheiter, 2009).

Students' source selections are thought to be guided by *information scent*. Information scent refers to learners' perceptions of the similarity between their cognitive representation of information need and the external representations or cues texts present (Kammerer et al., 2009). In other words, source selections are made when students

perceive a match between the information they are looking for and the information a source is purporting to include. A match perception is considered to be an indication of a strong information scent. In terms of the MD-TRACE model, the stronger a source's information scent, or perceived alignment with a student's task model, the greater the likelihood it will be selected.

The source selection literature has most commonly examined students accessing of texts from a search engine results page (e.g., Gerjets et al., 2011). Learners are considered to select texts from SERPs, based on sources' perceived topicality or relevance to task. Source topicality may be indicated to students through a variety of text-based external indices such as keywords or title (Kammerer et al., 2009). Rouet, Ros, Goumi, Macedo-Rouet, and Dinet (2011) identified two types of cues that may indicate topicality or relevance to students: surface cues and deep cues. When students make determinations of source topicality based on key-word matches or topographic indicators, such as bolding, they are attending to surface or superficial cues. When students make judgments about the type of information a source may offer or make inferences about a source based on document information presented they are said to be attending to deep cues. Surface cues allow for quicker and easier heuristic based decision making, whereas rendering relevance judgments based on deep semantic cues requires more effort (Rouet et al., 2011). However, deep semantic cues may be more efficacious for selecting sources to meet task demands (Rouet et al., 2011).

Studies have also varied modes of source presentation to examine its effects on source selection and subsequent multiple source use. Kammerer and Gerjets (2012a) examined the impact of search interface on students' source selections. Specifically, two

interfaces were compared—a list-like interface, simulating a search engine results page, and a graphic interface, wherein sources were presented as associated with particular ontological categories (e.g., disease diagnosis, or treatment). Kammerer and Gerjets (2012a) determined that when using a graphic interface, students selected more reliable sources more consistently (i.e., fewer commercial sites, more reliable sites). Likewise, Salmerón, Gil, Bråten, and Strømsø (2010) presented students with search results arranged either as a search engine results page (e.g., Google-like interface) or as a graphic interface mapping cause-effect relations (i.e., definition, causes, consequences, solution). Salmerón et al. (2010) found that arraigining sources via a graphic organization scheme led to improvements in multiple text comprehension and integration as the relations among texts were made explicit.

Beyond manipulating the arrangement of sources in a digital library, Le Bigot and Rouet (2007) examined differences in how search results appeared to students or the form of document information presented about each source prior to selection. Le Bigot and Rouet (2007) considered students' source selections when seven texts were presented in a hyperlinked menu either based on topic (i.e., title) or document information (i.e., author, date). Students were found to use source menus longer when document information was presented. Kammerer et al. (2009) manipulated how search results appeared to students by making some appear more topical or salient by bolding key words. The dwell time students devoted to source selection was examined. While students were found to consider almost all of the 30 available search results, they selected 32.91% of sites presented in a high-salience manner as compared to 25.22% of sites presented in a low salience manner.

Think-aloud data have been gathered to further explore the extent to which students consider criteria beyond topicality in source selection. However, students have been found to be limited in their selection of sources based on perceived quality (i.e., credibility, trustworthiness, reliability, or authority based cues, Brand-Gruwel, van Meeuwen, & van Gog, 2008; Gerjets et al., 2011). This may be in part because while information about topicality is readily indicated, information about source quality is more difficult to garner from search engine results, making source credibility a more difficult consideration to account for prior to source selection (Alexander & Tate, 1999; Britt & Gabrys, 2002).

Walraven et al. (2009) examined the criteria students provide for evaluating search results and selecting sources. Researchers determined six criteria students could use to evaluate search results: (a) source title or summary, (b) source type (e.g., pdf), (c) web-address, (d) rank in hit-list, (e) prior experience with the source, and (f) language of the source. Of these, source title, as an indicator of relevance, was almost exclusively the criteria students reported for evaluating search results (Walraven, et al., 2009). This is consistent with similar studies suggesting that students primarily select sources based on superficial relevance criteria rather than source reliability (e.g., Kuiper, Volman, & Terwel, 2005). Similarly, Gerjets, Kammerer, and Werner (2011) examined the types of justifications students introduced when selecting sources under two task conditions: either regular think-aloud instructions or instructions specifically prompting source evaluation. Students' justifications were coded into five categories, (a) topicality, (b) scope, (c) credibility, (d) design, and (e) up-to-datedness. Across conditions, students overwhelmingly evaluated sources based on topicality. However, credibility assessments

were also rendered when students were in the instructed evaluation think aloud condition.

In addition to think-aloud methodologies, students have also been asked to rate reasons for source selection. In a study with middle-school students, learners were asked to rate six source attributes to determine how *useful* each of these were in selecting sources (Braasch et al., 2009) and as indicators of potential source usefulness. The six attributes learners were asked to consider were: title, author, venue, date, and type of publication, and a 2-3-sentence source summary. Students were found to evaluate usefulness based on superficial factors and the most common differentiators for usefulness ratings were title and the description of source content; meanwhile, author and venue of publication, which could be used to identify source authoritativeness, were less common differentiators of source usefulness (Braasch et al., 2009).

Behavioral assessments of source selection have considered the number of sources or document types accessed. Kammerer and Gerjets (2012a) examined the frequency with which different types of sources (e.g., objective, subjective, commercial) were selected when search results were presented either in a list-like or tabular layout. Kammerer and Gerjets (2012a) determined that participants selected, on average, 46.32% of 18 possible search results. Objective search results were selected most often ($M=4.41$) and commercial search results were selected least often ($M=1.55$). Although there were differences in the number of different types of sources selected, the relation between number of sources used and response quality was not examined. This may be because of challenges in interpreting the significance of variable numbers of sources accessed. Stahl, Hynd, Britton, McNish, and Bosquet (1996) compared students' rates of access of 11 different documents to determine which occurred with greatest frequency. Stahl and

colleagues speculated that students' preferences for textbooks in source access may have been related to students' desire to obtain an overview of the relevant history or to use a *neutral* source (Stahl et al., 1996). However, the researchers noted that their interpretation was only speculative.

While there remains difficulty in interpreting number of sources accessed as an indicator of source selection, this remains an important consideration. First, examinations of number of sources accessed are rare because learners have most commonly been asked to access all documents within a particular source set (e.g., Bråten et al., 2009; Bråten et al., 2011). For example, even though students were operating under time constraints, Rouet, Favart, Britt, and Perfetti (1997) asked students to select each document in the library at least once. Second, considering the number of sources accessed is conceptually important as it is indicative of students' determinations of information need and recognition that their information needs have been met, in accordance with their task model. Stahl et al. (1996) found that the majority of situation model development, or creation of a coherent, cognitive representation of information presented across multiple texts, happened after students read the first two texts they accessed, with modest improvement in situation model development occurring after reading a third text. This suggests that accessing more than three sources has particular importance in possibly indicating limitations in situation model construction or in signifying an overly demanding task model. Pragmatically, the number of sources students select indicates the volume of information and range of viewpoints they have access to and are able to select from, compare, and draw on in composing a response. Beyond information need, an increasing number of sources accessed may indicate task

engagement. In the present study, number of sources accessed is be used as a metric of source selection.

Source Processing

Three approaches have been adopted in understanding students' source processing. In the most basic approach, learners have been asked to report the extent to which they relied on a variety of strategies when interacting with texts. Second, think-aloud methods have been used to access learners' multiple source processing. However, a limitation of such methods has been an interference with typical source use. Finally, time on texts has been used as an indicator of multiple text processing.

One of the more common methods for evaluating multiple source processing has been to ask students to complete strategy inventories following multiple source use. For example, Bråten and Strømsø (2006) asked students to complete a strategy use inventory that included three scales capturing the (a) memorization, (b) elaboration, and (c) monitoring strategies students engaged in when using multiple sources. In a later study, Bråten and Strømsø (2011) administered the Multiple-Text Strategy Inventory (MTSI), which included two dimensions: one addressing the accumulation of information from multiple sources and the other with items pertaining to cross-text elaboration, including students' comparing, contrasting, and integrating multiple texts. The first of these dimensions maps on to surface-level text processing or information aggregation while the second dimension considers deep-level processing or source evaluation and integration (Beizhuizen & Stoutjesdijk, 1999).

Similar aspects of strategy use have been examined when coding students' think-aloud utterances expressed during multiple source use. For example, Bråten and Strømsø

(2003) coded for five types of strategies in students' think-alouds, including multiple source use strategies focused on (a) memorization, (b) elaboration, (c) organization, (d) monitoring, and (e) evaluation of information found in multiple sources. Further, Bråten and Strømsø (2003) coded for think-aloud utterances related to students' strategies for text comprehension, including rereading, searching for clarifying information, using prior knowledge, searching for information in other sources, and skipping to read a later section.

In semi-structured, conversational interviews participants have been asked about the strategies they believed historians use as well as their own strategies for studying history (Hynd, Holschuh, & Hubbard, 2004). Prior to intervention, students have reported using low-level reading strategies including, re-reading, annotating, and highlighting; after instruction students have reported using higher-level strategies like comparing and contrasting texts, as well as using flashcards and organizational charts. This may suggest that learners have limitations in their strategic repertoire or limited knowledge of how strategies may be applied. Elsewhere, retrospective interviews have been used to examine students' metacognitive and monitoring strategies when interacting with multiple texts. For example, Stadtler and Bromme (2008) asked participants to report how much they knew about a topic, how well they believed themselves to have comprehended the information presented in multiple sources, and how much they believed they still needed to learn about the topic. Similarly, Strømsø and Bråten (2002) asked participants to report on their preparedness for an exam and if they have made any changes in study practices in relation to the examination - capturing students' assessment-focused monitoring.

Beyond having students report source use strategies, either prompted or not, behavioral metrics (i.e., time on text) have been used to examine students' source processing. Time on text is among the oldest examined indicators of single text reading and comprehension. Specifically, time on text has been found to correspond to text difficulty and structure (Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975), source type, students' engagement with text, and improved comprehension (Guthrie, Wigfield, Metsala, & Cox, 1999). Time on text has been also interpreted as indicating strategy use. For instance, learners have been found to devote more time to reading important rather than unimportant sentences (Duggan & Payne, 2011). Further, students have been found to devote more total gaze time to sentences that they could then accurately recall, as compared to inaccurately recalled sentences (Duggan & Payne, 2011). In studies of single text comprehension, where students have been asked to read regular versus refutational texts, learners have been found to fixate longer on contradictory statements (Ariasi & Mason, 2011). In a multiple text context, this suggests that students may spend longer on inconsistent or contradictory information presented across texts or process such conflicting information more deeply (Stadtler, Scharrer, Brummernhenrich, & Bromme, 2013).

When time on texts has been examined as an indicator of multiple source use, two primary processing strategies have been identified: *sampling* and *satisficing* (Reader & Payne, 2007). Sampling refers to readers first accessing each text in a multiple documents set, identifying the best source to meet task needs, and then devoting the majority of their source use time to that best text. This strategy is characterized by students first rapidly sequencing through sources with the goal of not specifically

learning about a topic, but rather, of first learning about the texts available. Students then devote the majority of study time to the best text and iteratively cycle to the next best source as their information needs change as a result of source use (Reader & Payne, 2007). Reader and Payne (2007, p. 269) describe the sampling strategy as a “commitment to choosing the best text available,” with a separation of source processing into an *exploration* phase, or the identification of the best source, and an *exploitation* phase, or the use of that source.

Satisficing is a strategy whereby students select a source they deem to be “good enough” and devote their source use time to that text. Satisficing is indicated by students’ study time being concentrated on the first source they access, assuming it meets a basic acceptability threshold. Rather than first learning about sources within a document set, as students do when engaging in sampling, when satisficing, readers are only motivated to learn about content in sources. Prior to engagement in source use, satisficing readers specify a level or standard of what would constitute an acceptable source, a potentially high threshold. If the first source accessed meets this threshold (i.e., threshold of satisfaction) students engage in linear reading (Foltz, 1996). However, when this threshold is unmet, students engage in skimming to ascertain threshold satisfaction, before moving to an alternate text.

Although sampling may appear to be the optimal strategy, both of these approaches to multiple source use are associated with particular benefits and drawbacks. For instance, while adopting a satisficing strategy may result in students devoting more time to lower quality texts, students also spend all of source use time on learning about a topic. Sampling is also contingent on students being able to judge attributes of texts with

relative speed and ease. More generally, Reader and Payne (2007) suggest that there are two standards to be met for sampling to be the preferred strategy. First, more than one text in a document set should be “good enough,” such that satisficing does not grant students access to all good sources. Second, there should be sufficient variation or “significant and perceptible differences,” between documents such that sampling may be necessary to gain a sense of each document (p. 270).

Notably, Reader and Payne (2007) suggest that students’ preferential allocation of time to various sources is indicative of subtle text judgments, beyond just those of source relevance. This is evidenced as, even when all texts in a document set are relevant, students differ in the amount of time they devote to using each source. For instance, students may draw conclusions about the amount of information they are able to learn from a source or a texts’ usefulness in meeting task goals in allocating source use time. Further, students may allocate study time based on perceived text comprehensibility or difficulty.

However, these time allocation strategies have not always been identifiable in students’ real-life source use. For instance, in a study where readers were provided with four texts about the human heart, students were found to visit each of the texts in the given document set for significant amount of times (Reader & Payne, 2007). On average, students were found to access 3.8 of four available texts, suggesting fairly comprehensive source engagement. Additionally, learners did not devote much time to any one source. Sampling, as a source use strategy, was used only rarely (Reader & Payne, 2007). Rather, students chose satisficing-like techniques, devoting considerable time to each source accessed, in sequence. Wilkinson, Reader, and Payne (2012) corroborated these

findings, determining that 76.7% of students could be classified as satisficers, 6.7% as samplers, and 16.7% of students demonstrated browsing behaviors that could not be classified.

Reader and Payne (2007) explain the amount of time students devote to source access by suggesting that each of the texts presented to students included unique information that learners would want to access. Other studies have found learners to open web texts only briefly (e.g., less than ten seconds), although 10% of sites are accessed more extensively (i.e., for more than two minutes, Weinreich, Obendorf, Herder, & Mayer, 2008). Duggan and Payne (2011) suggest that students may begin multiple source use by engaging in a sampling strategy, accessing texts briefly, before adopting a satisficing approach. More generally, students' source access and time allocation may only be systematic to the extent that students do not revisit sources until each text has been visited at least once (Reader & Payne, 2007).

Once a text is accessed, some have suggested that source processing may not be uniform. While students may access and use each text in a set, for more favored texts, time is devoted to reading, while for less favored texts, time is spent on skimming (Reader & Payne, 2007). Duggan and Payne (2011) used eye-tracking data to corroborate the prevalence of students' engagement in skimming or scanning of each text accessed while adopting a satisficing strategy. Although students did engage in linear reading of important text passages, once text usefulness dropped below a particular threshold, learners were, indeed, much more likely to skim or rapidly process texts in a non-linear fashion. In particular, skimming may be effective because it distributes attention throughout a source and may help students better attend to the macro-structure

of a text, supporting comprehension to a greater extent (Duggan & Payne, 2011).

Based on findings that students generally engage in satisficing or otherwise indeterminate behaviors and that source processing within-text alternates between skimming and reading (Duggan & Payne, 2011), in the present analyses students' time on texts is collapsed and used as a general measure of processing (e.g., Reader & Payne, 2007; Wilkinson et al., 2012). Rouet, Vidal-Abaraca, Bert-Erboul, and Millogo (2001) provide additional evidence for considering students' time on texts globally. In a study of students' source use when responding to high level questions (HLQs), requiring multiple source integration, versus low level questions (LLQs), such as information location tasks, Rouet et al. (2011) found evidence to suggest that when responding to HLQs, the source use process was uniform and distinct from the source use pattern students exhibited when presented with LLQs. Students responding to HLQs were generally found to access more paragraphs and to more readily engage in an *integrative search pattern* or systematically visit paragraphs in texts. Rouet et al. (2011) term students' source use *review and integrate*, suggesting that after constructing an initial cognitive model of a text, students made successive passes through a source modifying their understanding. This processing pattern seems to indicate that a global measure of time on texts is appropriate to summarize students' depth of text processing and engagement.

Notably, the satisficing and sampling strategies identified by Reader and Payne (2007) emerged when students identified information while respond to multiple-choice questions within a restricted period of time (i.e., seven minutes or 15 minutes). Even when responding to such LLQs, students were found to evidence a generally uniform

pattern of source use (i.e., satisficing). Unlike in the Reader and Payne (2007) study, when students' time on texts is not restricted by task parameters, the amount of time learners freely devote to source use may be an even more effective metric of overall text processing and task engagement.

A single metric approach to assessing source processing was used by Cerdán, Vidal-Abarca, and colleagues who used the Read&Answer software to track students' reading time (Cerdán & Vidal-Abarca, 2008; Cerdán, Vidal-Abarca, Martinez, Gilabert, & Gil, 2009; Vidal-Abarca et al., 2011). Cerdan and Vidal-Abaraca (2008) identified three measures of source processing computed based on reading time. The first, total time on text, was considered as a summative measure of source use. The second measure was a totaling of the amount of time students devoted to processing relevant and irrelevant information. The third metric, considered to be a measure of source integration, was a totaling of the time students devoted to reading relevant, non-consecutive sections of text, as this was thought to suggest information assembly or integration. Across these three measures, students' source use was timed globally, across texts, rather than in a source-specific manner.

While Cerdán and colleagues used three metrics, only the first measure, total time on texts, is used in the present study. In the Cerdán and Vidal-Abarca (2008) study, students were expected to include specific idea units in composing their responses to either intra-textual or inter-textual questions. The specific idea units to be included were coded as relevant while all other idea units in texts were coded as irrelevant. Further, not all texts included in the document set were relevant for responding to each of the questions. In the present study, there are no specific ideas that are needed to be included

in students' written responses and any of the texts included in the source library are potentially relevant in composing a response. For these reasons, passage relevance based indices of time on texts are not examined. Additionally, using only overall total time on texts as a metric of processing allows for unobtrusive data collection, whereas software such as Read&Answer requires partial masking and unmasking of relevant and irrelevant sections of text to time students' processing.

Using a global metric of time on texts may be particularly appropriate within a multiple documents framework (Rouet & Britt, 2011). When students are presented with an array of texts to choose from, they can be more selective and if unsatisfied with a particular source readily move to an alternate text. Because there is a greater volume of information available, students may feel more comfortable skimming or processing selectively. As a result students' time spent on source use may be more variable in a multiple source use context and less directly related to the level of processing adopted, than when learners use a single texts (Rouet & Britt, 2011), justifying the use of a single measure.

Source Evaluation

Source evaluations can be defined as students' judgments of or conclusions about sources or the information within them based on "available or accessible information about the source" (Bråten et al., 2009, p. 6). Broadly speaking, two categories of source evaluations can be identified. The first are epistemic source evaluations and the second are non-epistemic source evaluation. Epistemic dimensions of source evaluation are those that are aligned with students' epistemic beliefs (i.e., concerned with knowledge establishment) or based on considerations of document information (e.g., author

credentials or expertise). Epistemic based evaluations include students' judgments of sources based on quality-related criteria. In particular, students have been asked to evaluate sources along trustworthiness (Bråten et al., 2009; Bråten et al., 2011), reliability (Wiley et al., 2009), and authoritativeness (Mason et al., 2010a, 2010b) criteria. Non-epistemic dimensions of source evaluation are based on criteria not concerned with knowledge establishment and not reflective of quality-based considerations. Rather, non-epistemic dimensions of source evaluation are focused on factors related to task-relevance or superficial source features (e.g., source presentation or formatting).

Three primary methods of assessment have been used in gauging students' source evaluations. First, questionnaires have been administered, typically ahead of source use or in place of a source use task, to survey students regarding their beliefs and behaviors when evaluating sources (e.g., Flanagin & Metzger, 2000). Second, in-process measures of source use have been used; methods such as think-alouds capture students' evaluative utterances while completing tasks (e.g., Mason et al., 2010b). Finally, source evaluations have been examined using a dimensional approach, wherein students have been asked to evaluate texts according to some researcher specified criteria. Dimensional evaluations have been elicited both during (Stadtler & Bromme, 2007, 2008) and subsequent to source use (e.g., Bråten et al., 2009). Students have been asked to render judgments either by rating sources along various dimensions or by ranking sources, relative to one-another, on a particular characteristic (e.g., reliability, Wiley et al., 2009). Following dimensional judgments of sources, students have been asked to justify their evaluations to varying extents.

Learners have been surveyed with regard to their general attitudes toward source

evaluation and behaviors when judging sources (e.g., Hargittai et al., 2010; Purdy, 2012). For instance, Metzger, Flanagin, and Zwarun (2003) asked 356 undergraduate students to report the frequency with which they engaged in nine different verification strategies when encountering information online. The verification strategies students were asked to endorse included: (a) checking information currency, (b) checking information completeness, (c) determining whether information constituted fact or opinion, (d) validating information across other sources, as well as considering (e) the author, (f) the author's objectives, and (g) credentials. On average, students reported checking information currency and completeness *occasionally* or *often*, however, they checked author, author objectives, and credentials only *rarely* or *occasionally* (Metzger et al., 2003).

Despite their use, particularly in exploratory studies of source evaluation (e.g., Metzger et al., 2003), survey methods have been critiqued for their reliance on self-reported behaviors, rather than assessment of evaluations in context (Metzger, 2005). Further, Purdy (2012) suggests that students' reports of engagement in evaluative behaviors may be susceptible to a desirability bias; students may report engaging in verification strategies that they do not in fact use when completing academic tasks. At the undergraduate level, it may be expected that students are able to report the source evaluation strategies they *ought to* engage in but there is a need to better identify the actual evaluative practices used by students when interacting with texts (Metzger, 2007).

To address these limitations, two general approaches have been adopted to examine students' evaluative behaviors during task engagement. First, students have been asked to think-aloud while completing tasks and their evaluative utterances have

been recorded. Second, source evaluations have been elicited during task completion by asking learners to rate texts along specific dimensions while engaged in source use.

Mason et al. (2010b) examined students' spontaneous source evaluations reported while thinking-aloud during Internet research. Students' evaluative utterances were coded as corresponding to particular dimensions of students' epistemic beliefs (i.e., beliefs about knowledge and knowing, Hofer & Pintrich, 1997). Two epistemic belief dimensions in particular, those of students' beliefs about the *source of knowledge* and beliefs regarding *justifications for knowledge*, were, respectively, considered to be particularly pertinent to the evaluation of sources and the information within them. Students' evaluative utterances, consistent with the source of knowledge dimension of epistemic beliefs, were coded at three levels (Mason et al., 2010b). Learners could have evaluated (a) the *popularity of a source*, (b) its *authoritativeness or expertise*, or the (c) *scientific nature of the source*. These dimensions were progressive in sophistication, with judging sources based on their scientific nature considered to be the most sophisticated evaluative dimension. Likewise, in evaluating the information within a source, consistent with the justifications for knowledge dimension of epistemic beliefs, students' utterances were coded into three categories, progressing in sophistication. Learners could have determined that information was (a) *unable to be evaluated*, (b) *agreed or disagreed with their own beliefs*, or was (c) *based on scientific evidence*.

While Mason et al. (2010b) only considered students' epistemic source evaluation, evidence suggests that non-epistemic source evaluations are crucial to consider as well. In a study by Walraven et al. (2009), students' source evaluations were captured by having high-school students first think-aloud during an Internet search task

and later participate in a follow-up focus group interview about their source and information evaluations. While researchers identified an exhaustive, twenty-nine potential criteria along which students could evaluate search results, sources, or the information within them (Walraven et al., 2009), students reported using only a minority of these in judging the texts they encountered.

During the think-alouds, sources, as a whole, were evaluated very rarely; in cases that they were, the criteria applied most commonly was *speed*, or the amount of time a page required to load. In evaluating information in sources, the vast majority of judgments concerned the *connection of information* to the target task or information relevance. Other criteria cited by students included the *language* of the site and the *amount* of information offered; all three of these (i.e., information connection, language, and amount of information) were classified under the criteria of *site usability*. Language referred to whether the site used domain-specific language or had spelling or grammatical mistakes. Amount of information referred to the volume of information a site offered and whether a source provided all of the information students needed in completing a task or whether additional information was required. In students' evaluations of both sources and the information within them, non-epistemic criteria (e.g., speed, relevance, amount of information) were most commonly reported.

The evaluative criteria most often cited by students in the study by Walraven et al. (2009), connection of information to task or information relevance, is the none-epistemic dimension of source evaluation that has received most attention (e.g., Anmarkrud et al., 2013). Evaluations of relevance can be defined as students' determinations that a source serves an instrumental purpose in meeting their goals for source use or as a judgment of

task-based importance (Rouet & Britt, 2011). Relevance judgments have further been defined as being either positive (i.e., related to task) or negative (i.e., unrelated to task) and as pertaining to texts as a whole (i.e., source relevance) or to specific content (i.e., information relevance, Anmarkrud et al., 2013).

Determining relevance is considered to be an essential evaluative process for students to engage in when interacting with multiple texts (e.g., Rouet & Britt, 2011). Anmarkrud et al. (2013) suggest that students' relevance judgments may serve as a precursor to their engagement in source evaluations along more epistemic dimensions. Before deciding to process a text and engage in the effortful evaluation of its trustworthiness, students may first determine its potential utility in helping them to meet their goals for source use (Anmarkrud et al., 2013; Rouet & Britt, 2011). Further, such judgments may aid students in linking related information, identified as relevant, across texts. Given the role that relevance judgments play in students' subsequent source evaluations and in text comprehension overall, it is an encouraging finding that students, at least at the undergraduate level, have been found to be more skilled and more frequent in making judgments of relevance (Anmarkrud et al., 2013) compared to the rate at which they engage in more effortful evaluations along epistemic criteria.

A challenge in examining students' in-process source judgments is determining whether or not students ought to be explicitly instructed to engage in source evaluation. On the one hand, when not explicitly instructed to evaluate texts, students have been found to evaluate sources only to a limited extent (e.g., Walraven et al., 2009). On the other hand, explicitly instructing students to engage in source evaluation may interfere with typical processing and may distort the number and types of evaluations offered. In

one study, Gerjets et al. (2011) manipulated this specific consideration by directing students either to simply think aloud (i.e., spontaneous evaluation condition) or further asking them to evaluate sources as they did so (i.e., instructed evaluation condition). Gerjets et al. (2011) found that while students spontaneously evaluated sources to a limited extent, as expected, they did so more frequently when explicitly instructed. Specifically Gerjets et al. (2011) found that when spontaneously evaluating sources online, students most often judged webpages according to site credibility and scope; when students were explicitly instructed to evaluate websites while thinking aloud, they considered credibility criteria to a much greater extent and also evaluated website design (Gerjets et al., 2011). A number of studies have determined that students provide much more frequent evaluations and evaluations along differing criteria when explicitly instructed to evaluate sources, as compared to when spontaneous evaluations are examined (e.g., Gerjets et al., 2011; Stadtler & Bromme, 2007, 2008).

Students' limitations in spontaneously generating source judgments are considered to be suggestive of challenges with source evaluation more generally. The particular source evaluation mechanism students are considered to be challenged by is *sourcing* (e.g., Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012; Wiley et al., 2009; Wineburg, 1991). Sourcing is defined by Wineburg as, "looking first to the source or attribution of the document" (1991, p. 79) and further includes attention to any document related information that may aid in source evaluation (Britt & Aglinskas, 2002). Sourcing often involves considering identifying document information (e.g., author, publisher) associated with a source's origin or purpose for being written. Difficulties in sourcing may arise from students being limited in attending to source

information when judging texts (Britt & Aglinskias, 2002) or in failing to properly make use of appropriate document information in evaluation sources (Bråten et al., 2011).

For instance, Bråten, Strømsø, and Salmerón (2011) found low knowledge students to be less discriminant in using different types of document information to evaluate sources and to attend to errant trustworthiness indicators (e.g., date). In recognition of such difficulties, Stadtler and Bromme (2007, 2008) introduced a software tool, *met.a.ware*, to promote students' engagement in the epistemic and metacognitive processes of source evaluation. The *meta.a.ware* program asks students to systematically log each piece of information pertinent to their query in an online system. In logging each piece of information, students are also asked to record the document it originated from, encouraging sourcing. Researchers also introduced an evaluative condition wherein students were additionally asked to rate three statements pertaining to each recorded document. Specifically, students were asked to rate (a) *author's expertise* and (b) the potential *bias or commercial interest* of a source, as well as their (c) *confidence in the information* logged.

Adding to the discussion of whether or not students' source evaluations ought be prompted, relative to the control group, students in the evaluative condition performed significantly better on a follow-up sourcing test and included a higher percentage of correctly sourced arguments in their essays (Stadtler & Bromme, 2007). It seems that prompting source evaluation results not only in increased evaluative behaviors but also in general learning improvements (Stadtler & Bromme, 2007, Wiley et al., 2009). Although a think aloud methodology was not employed in the Stadtler and Bromme study (2007),

students' source evaluations were able to be assessed during task completion by collecting evaluative data in situ, within the met.a.ware interface.

Although Stadtler and Bromme (2007, 2008) did not examine students' actual source evaluations in the met.a.ware system, electing to only compare learning outcomes for students in the evaluative and control conditions, prompting source evaluation while students are engaged in source use seems a promising avenue for collecting in-process reasoning about texts. Further, embedding source evaluations into a digital system (e.g., Stadtler & Bromme, 2007, 2008) allows for the collection of in-process evaluation data without the intrusiveness and cognitive demands that are considered to be limitations of think-aloud methodologies (Mason et al., 2010b). While Stadtler and Bromme (2007, 2008) asked students to rate texts during the course of source use, the majority of studies asking students to judge texts along specific dimensions have considered retrospective ratings, made following task completion. As with in-process measures, these retrospective examinations of source evaluation have focused on students' source judgments rendered based on document information.

In examinations of source evaluation, students have most often been asked to judge sources either according to predefined criteria (i.e., rate, Bråten et al., 2009) or in relation to one another (i.e., rank, Wiley et al., 2009) and to justify these evaluations to varying extents. For example, Bråten, Strømsø, and Britt (2009) examined students' trustworthiness ratings after reading seven texts on the topic of climate change. Students were first asked to rate the trustworthiness of the information in each source; then, they were asked to report the extent to which they based each trustworthiness ratings on one of six characteristics: (a) author, (b) publisher, (c) source type, (d) content, (e) personal

opinion, or (f) date of publication. The two sources students found to be most trustworthy were from respected, research organizations in Norway (i.e., CICERO, NPCA); these sources were also deemed by researchers to be the two most trustworthy documents in the set. In terms of justifying their trustworthiness ratings, students most often cited the content of the source, followed by considerations of author and document type. Such a methodology, asking students to render trustworthiness ratings and to justify these based on document information, has been adopted in other studies as well (e.g., Bråten et al., 2011; Strømsø et al., 2011).

In addition to examining how students rate sources along trustworthiness dimensions, learners have been asked to rank sources as well as to consider additional dimensions of source evaluation. Rouet, Britt, Mason, and Perfetti (1996) asked students to rank sources with regard to *trustworthiness* (i.e., “to the extent you trust what the author says”) and *usefulness* (i.e., “to the extent they helped you build up an informed opinion during the study period”). Students were also asked to justify each source ranking produced. For trustworthiness rankings, four justification criteria were identified. These were: (a) author-related justifications, (b) document type-based justifications (e.g., textbook, newspaper), (c) content-based justifications, referencing specific information in the source, and (d) opinion-based justifications, stemming from students’ personal views about a described issue. Usefulness rankings, referring to the pragmatic value of texts for meeting task goals, were not justified. Rouet et al. (1996) found students’ trustworthiness and usefulness rankings to differ for different source types, but to be consistent across these. For instance, students ranked the textbook source as both the most trustworthy and the most useful in the document set.

There are a number of limitations in the efficacy of collecting post-hoc source evaluations. First, these are retrospective and therefore may draw not only on students' evaluations of the specific text targeted, but also on students' judgments of the other texts, the task, or their performance (Goldman et al., 2012). Few studies provide students with explicit instruction regarding which target or referent to consider when evaluating sources. As an exception, Wolfe and Goldman (2005) asked students to evaluate the similarity and difference in arguments between two contradictory texts. However, in most studies participants have not been explicitly directed to evaluate any particular target. It is also unclear whether evaluations reported constitute students' initial judgments, made during source use, or post-hoc evaluations, rendered only when prompted to do so. Further, post-hoc data collection requires students to recall source information, or remember each text and correctly associate the information within it; students have been found to be imperfect in their memory for sources (Strømsø, Bråten, & Britt, 2010).

In the present study, as in prior work (Rouet et al., 1996; Rouet et al., 1997), students were asked to rate sources along two dimensions, one epistemic (i.e., trustworthiness) and one non-epistemic (i.e., usefulness). Usefulness judgments reflected students' evaluations of texts in reference to task demands. Anmarkrud et al. (2013) suggest that usefulness evaluations are a precursor to further text processing. Trustworthiness judgments are considered to be based on evaluations of document information (e.g., source type, author, publisher). Bråten et al. (2011) describe trustworthiness evaluations as an *essential aspect* of multiple source use, particularly online. Jointly, these two evaluative dimensions serve to reduce the volume of

information load students need to contend with (Strømsø et al., 2010). Students may more easily dismiss information from sources deemed low in utility or in trustworthiness. Further, judgments along these dimensions may aid students in contextualizing and, therefore, better processing text content. Usefulness judgments may serve as a filter, helping students to process sources in task-referential ways. Trustworthiness judgments are thought to reflect students' attention to document information. Therefore, these imply students considering text purpose and the context of how a source was created and elaborating information based on document information (Strømsø et al., 2010; Wineburg, 1991). Given the dependency between students' evaluations of trustworthiness and considerations of document information (e.g., Britt & Aglinskias, 2002; Britt et al., 1999; Perfetti et al., 1999; Strømsø et al., 2010), an additional evaluation metric was examined in this study: whether or not students elected to access document information for each source they used. Accessing of document information has been considered to underlie expert-level multiple source use (Wineburg, 1991). Specifically, sourcing is considered to be associated with text contextualization and with multiple source corroboration, or the reconciliation of discrepant information across texts (Wineburg, 1991) and to improve multiple text comprehension (Strømsø et al., 2010).

Students were asked to rate each source used along an additional dimension, interestingness. Despite the important role of engagement in multiple source use (Perfetti et al., 1996), interest has been examined in the multiple source use literature only to a limited extent (Hidi & Baird, 1986).

Interest has been distinguished as constituting either a sustained personal preference toward a particular topic or domain (i.e., individual interest) or as arising from

the context surrounding a particular task or the “interestingness” of a situation (i.e., situational interest; Alexander, 1997; Hidi & Baird, 1986; Renninger, Hidi, & Krapp, 1992). When examined in the MSU literature, interest has been most commonly assessed at the topic level, prior to students’ task engagement, rather than as arising situationally (e.g., Bråten, Gil, Strømsø, & Vidal-Abarca, 2010; Strømsø, Bråten, & Britt, 2010).

A form of interest, *text-based interest* or *text interestingness*, has been introduced to reflect how aspects of topic or text may engage students’ situational interest during source processing (Hidi & Anderson, 1992). Kintsch (1980) has suggested that interest may arise during text processing in two ways: through engaging or surprising events occurring in text (i.e., *emotional interest*), more common in narrative compositions, and through elements in text connecting with students’ prior knowledge (i.e., *cognitive interest*). Likewise, Hidi and Baird (1988) have distinguished between texts inspiring knowledge-triggered interest and value-triggered interest. Knowledge triggered interest develops when information in text connects to students’ prior conceptions, experiences, or schema. Value triggered interest arises when text content relates to readers’ “values, desires, and preferences” (Hidi & Baird, 1988, p. 469).

Although text-based interest has been associated with text recall and reading comprehension (Hidi, 2001; Hidi & Baird, 1986; Krapp, 1999; Schiefele, 1999), Hidi and Baird (1986) have discussed the difficulties of identifying which facets of expository text may give rise to text-based interest. While Hidi and Baird (1988) identified elements associated with expository text interestingness for elementary school students (e.g., novelty, salient elaborations of main points, embedded “need for resolution”), work has been more limited in identifying facets of texts that may inspire interest in older learners.

Further, some efforts to make sources more engaging have compromised learning. For instance, the inclusion of *seductive details* (i.e., irrelevant or extraneous information included in texts to increase interest) in expository writing has been found to interrupt text coherence and to interfere with situation model construction (Garner, Brown, Sanders, & Menke, 1992; Garner, Gillingham, & White, 1989; Harp & Mayer, 1998). Likewise, Sanchez and Wiley (2006) found that for undergraduate readers, the inclusion of images alongside expository texts both increased interest and had a seductive effect resulting in lower comprehension, particularly for students with low working memory.

Text interestingness has not yet been examined, to my knowledge, in a multiple text context. Nonetheless, text interestingness may be a particularly important factor in students' multiple source use for a number of reasons. First, when students are presented with multiple text options, interestingness may be a dimension, alongside source trustworthiness and usefulness, that learners employ in discriminating between sources and in narrowing the scope of available information they are required to consider. Further, when students are presented with multiple texts, varying in interestingness, they may be more liberal in navigating among text options and may more readily reject sources low in interest, without concern for reducing the volume of information they have access to. Researchers have suggested that interest serves as a motivational variable, particularly when students' experience challenges in text processing, or as a "mental resources for learning" (Boscolo & Mason, 2003; Mason & Boscolo, 2004). In multiple text contexts, presenting greater cognitive complexity than single text tasks (Bråten & Strømsø, 2006), interest may be especially important as a driver of students' engagement with texts and as a motivational reserve available when challenges arise.

Motivational benefits of interest extend beyond simply greater persistence and may enable students to gather a greater volume of information, across texts. Additionally, interest as a motivational factor has been found to allow for deeper text engagement, such as that resulting in conceptual change (Andre & Windschitl, 2003; Dole & Sinatra, 1998). Hidi (2001; 2006) makes the case that interest engages a unique type of cognitive processing, not available absent interest. This type of interest-based processing is characterized by an exploratory orientation toward content, deeper and more elaborative engagement with texts, and a propensity for developing connections between information and prior experience (Wade, Buxton, & Kelly, 1999). As multiple source use is typically required for complex tasks, the type of cognitive engagement interest supports may be particularly facilitative to performance.

In this study, to capture text-based interest, students were asked to rate the interestingness of each source accessed. Students' average rating of interestingness, across all sources accessed, was considered a measure of learners' motivational and cognitive engagement during text processing. In prior work, Mason and Boscolo (2004), rather than assessing topic interest, likewise, asked students to rate text interestingness during source use.

Students were asked to access document information and rate texts during source use rather than retrospectively, with these functionalities embedding into the online interface (e.g., Stadtler & Bromme, 2007, 2008).

Individual Difference Factors in Multiple Source Use

In addition to considering the above described source use behaviors (i.e., source selection, source processing, and source evaluation), three learner characteristics were examined as impacting multiple source use. The next section describes the role of prior knowledge, attitudes, and domain general source evaluation behaviors in students' multiple source use.

Prior Knowledge

More than any other learner characteristic, prior knowledge has been most extensively examined in relation to both single and multiple text processing (Alexander et al., 1994; Alexander et al., 1991; Kerstetter & Cho, 2004). Indeed, prior knowledge has been shown to effect all phases of the multiple source use process, including source selection (e.g., Potelle & Rouet, 2003), source processing (e.g., Rouet, 2003), and source evaluation (e.g., Bråten et al., 2011).

At the same time, the effects of prior knowledge have been most commonly investigated when students access and navigate *hypertexts* (e.g., Potelle & Rouet, 2003; Salmerón, Kintsch, & Cañas, 2006; Salmerón et al., 2010) or information nodes directly linked to one another (Rouet, Levonen, Dillon, & Spiro, 1996). In studies of multiple source use, when texts are not explicitly linked, prior knowledge has most commonly been explored as a control variable (e.g., Strømsø et al., 2010; Strømsø, Bråten, & Samuelstuen, 2008) despite evidence suggesting that in such multiple text contexts, where sources are not explicitly connected, prior knowledge may matter all the more (Britt & Aglinskias, 2002; Rouet et al., 1997).

In the sections to follow, the role of prior knowledge in source selection, processing, and evaluation is discussed, as impacting ultimate task performance.

Prior knowledge and source selection. In examining prior knowledge and source selection, research has considered the extent to which various source presentations formats may benefit students with differing levels of prior knowledge. Generally speaking, sources connected or explicitly linked to one another have been found to be preferable to alternate listings (e.g., alphabetic, Möller & Müller-Kalthoff, 2000; Potelle & Rouet, 2003; Shapiro, 1999). For instance, Potelle and Rouet (2003) examined the role of prior knowledge in students' accessing of sources presented via three types of representation systems. Sources were displayed to students via a hierarchical map, a network map, or an alphabetic list. The hierarchical map linked texts progressively, from those addressing most general to most specific content. The network map connected text content (i.e., main ideas) with semantic links; for instance, the network map linked texts that identified causes with those presenting effects. No explicit connections between texts were offered in the alphabetic menu presentation. It was determined that while for high knowledge students' performance did not differ across the three source presentation conditions, for low knowledge students, the hierarchical representation structure was preferable to the other two conditions (Potelle & Rouet, 2003).

These findings suggest an interaction between learners' prior knowledge and source selection. While high knowledge students may possess the requisite prior knowledge and skills to conceptualized multiple sources in relation to one another, students with low prior knowledge benefit from having connections between multiple texts made explicit. Low knowledge learners' preference for source presentation in structured ways has been found across studies (McDonald & Stevenson, 1998; Möller & Müller-Kalthoff, 2000; Potelle & Rouet, 2003; Shapiro, 1999). However, an important

qualifier to these findings comes from a study by Hoofman and van Oostendorp (1999). Researchers presented texts to students either through a conceptual map, identifying relations between texts, or alphabetically. They found that for low knowledge students, concept maps may be too complex a mode of source presentation; they suggest that a more basic content listing may be preferable.

Importantly, across studies, students' prior knowledge in relation to sources selected *per se* has not been examined. Rather, learners' prior knowledge and mode of text presentation have been considered as predictors of task performance (e.g., Potelle & Rouet, 2003). In the present study, the relation between prior knowledge and source selection (i.e., number of sources accessed) is directly considered, as are their joint effects on response quality.

Work by Salmerón, Kintch, and Kintch (2010) offers some initial insights into the relation between prior knowledge and source selection. Salmerón et al. examined the criteria students used to select the order in which they accessed sections of a hypertext (2010). Specifically, three criteria for source selection were identified: *coherence*, *interest*, and *default screen position*. Coherence refers to students' decisions to select texts most similar to sections previously read. Interest refers to students' decisions to select most interesting texts first, while delaying selecting less interesting sections of text. Finally, the default screen position criterion was invoked by students choosing sources depending on their location or order of presentation in a hyperlinked menu.

Of these three strategies, the coherence strategy was considered to be the most cognitively demanding, requiring students to infer conceptual relations between texts (Salmerón, Kintsch, & Canas, 2006; Salmerón et al., 2010). Low knowledge students

benefitted most from adopting coherence criteria in selecting hypertexts; however, there were no performance differences across selection criteria for students with average or high levels of prior knowledge (Salmerón et al., 2006; Salmerón et al., 2010). This suggests that low knowledge students may be the ones most in need of support to conceptually connect multiple texts, whereas students with intermediate to high levels of prior knowledge do so quite readily (Potelle & Rouet, 2003; Salmerón et al., 2006). Relating multiple texts to one another may be fostered either by menu arrangement or by students adopting an orientation toward considering text selections in this way.

Although hypertext selection criteria were found to be differentially adaptive for students with varied levels of prior knowledge, there has been limited work associating prior knowledge with the type of hyperlink selection criteria students adopt. This may be because there are challenges associated with definitively classifying students into a particular group (e.g., coherence-based selectors) or because learners have been found to use a combination of criteria in navigating hypertexts (Salmerón et al., 2006). In the present study, the role of prior knowledge is examined as it relates to a more global, behavioral measure of source selection, the number of sources students elect to access. Initial findings suggest that this is a promising area for investigation; within hypertext contexts, McDonald and Stevenson (1998) found knowledgeable students to access more nodes than non-knowledgeable students. Indeed, high knowledge students' abilities to identify coherence between texts may facilitate their accessing of a higher number of sources.

Prior knowledge and text processing. Prior knowledge is generally considered to underlie all text processing (Fincher-Kiefer, 2005; Langer, 1982; Soederberg Miller,

2001). In Kintsch's Construction Integration Model of comprehension (Kintsch & van Dijk, 1978) the situation model students generate in comprehending texts is not a direct representation of information in a source. Rather, the situation model constructed moves beyond the text-based model, representing text-direct propositions, to integrate information in a source with students' prior knowledge. Potelle and Rouet (2003) make the case for the universality of this integration process as all texts are written assuming that readers have prior knowledge, not explicitly presented in texts, that they can draw on in comprehending and generating inferences while processing sources.

Amadiou, Tricot, and Mariné (2009) refer to prior knowledge as a cognitive resource that can both aid students in text processing and reduce the processing demands associated with multiple source use. Specifically, prior knowledge may help students in managing three types of text and task related information: (a) *intrinsic cognitive load*, or the volume of information inherent in a text-based task, (b) *extrinsic cognitive load*, or instructional or environmental constraints that may make the task more challenging, and (c) *germane cognitive load*, or the cognitive effort required to learn from texts.

In addition to its facilitative features, prior knowledge may help learners in navigating challenges associated with multiple source use. Among these is *disorientation*, defined as a difficulty in identifying one's position among texts and knowing subsequent sources to visit (Amadiou et al., 2009). Two types of disorientation have been considered: *structural disorientation*, referring to challenges with multiple text navigation, and *conceptual disorientation*, referring to difficulties with the linking or integration of multiple sources, interfering with situation model construction. When disorientation is experienced it contributes to excess cognitive load and may impede

successful text processing (Amadiou et al., 2009). The role of prior knowledge in preventing disorientation is supported by findings that students with higher levels of prior knowledge revisit texts previously read less frequently than do low knowledge students (Symons & Pressley, 1993).

Schema-based theories of text comprehension have also offered insights into the mechanisms by which prior knowledge (i.e., conceptual information stored in long-term memory) may support text processing (Symons & Pressley, 1993). Indeed, prior knowledge may help in identifying relevant information in texts, drawing appropriate inferences or conclusions from sources, understanding when texts are confusing or ambiguous, integrating texts, and in text recall (Symons & Pressley, 1993).

Selecting relevant information. Prior knowledge schema may help learners to select relevant information from texts that is knowledge consistent. It may be the case that prior knowledge serves to direct students' attention to important and relevant information when processing texts or that new information encountered is integrated into gaps within prior schema (Symons & Pressley, 1993). Prior knowledge may also aid students in determining what is task-relevant and selecting such information (Rouet, 2003).

Inference generation. Prior knowledge is considered to also support inference generation or students' connecting information in text with prior knowledge (Rouet, 2003). In studies of single text processing, readers with higher levels of prior knowledge have been found to compensate for a lack of explicit information presented in text (McNamara et al., 1996) and to use inferencing to cope effectively with low coherence sources. McNamara et al. (1996) found that while students with low levels of prior

knowledge better comprehended texts where information presented was explicit (i.e., limited in prior knowledge demands), high knowledge students benefited from less coherent texts that required them to activate their prior knowledge to generate inferences and therefore supported deeper text processing. Generally, students with more prior knowledge to connect to textual information more readily generate inferences in constructing a situation model.

Reduced cognitive load. Prior knowledge may also reduce the cognitive load associated with text processing (Guthrie, 1998; Symons & Pressley, 1993). The memory related benefits of prior knowledge offer particular benefits when solving multiple text tasks. This is because in order to respond to complex problems (e.g., multiple source tasks), students have to hold intermittent information garnered from texts in memory (Guthrie, 1998). Having high levels of prior knowledge supports this intermediate step, as students are better able to integrate new information into existing schema, thereby retaining it in memory with greater ease.

Reduced cognitive load, resulting from easier surface-level text processing, may enable engagement in higher-level strategic processing, such as generating inferences or integrating multiple texts. Students with reduced cognitive load may further be better able to attend to macro-level structural features of texts, supporting source comprehension and integration. Sweller (1994) suggests simply that complex multiple text problems may be too cognitively demanding for students low in prior knowledge.

Information integration. In addition to reduced cognitive load, prior knowledge confers a *storage advantage* whereby new information is better integrated and retained in existing schema (Symons & Pressley, 1993). Potelle and Rouet (2003) argue that when

students build a situation model of text comprehension, they engage in the *reorganization* and *restructuring* of text-based information according to prior knowledge schema and integrate new information into existing schematic understandings. Better integration results in better storage of information in students' long-term memory and improved recall of situation model elements (Amadiou et al., 2009; McNamara et al., 1996; Rouet, 2002; Wiley, 2005). In multiple source use contexts, prior knowledge may facilitate the integration of relevant information both across text sections and across texts (Afflerbach, 1990).

Text organization. High-knowledge learners may better integrate texts because they are better able to organize information within and across sources during processing (Amadiou et al., 2009; Rouet, 2003). Prior knowledge may promote “incidental learning” of document structure during source engagement (Rouet, 2003, p. 14), as demonstrated by studies in which students with greater topic familiarity recall a greater number of text headings. Additionally, students with higher levels of prior knowledge have been found to devote a greater portion of source use time to examining source presentation categories and organizational indicators (e.g., table of contents, Dreher & Guthrie, 1990).

Prior knowledge may specifically help students with organizing or categorizing multiple texts in relation to one another through the identification of task-relevant source categories (Potelle & Rouet, 2003; Salmerón, Kintsch, & Cañas, 2006). For example, students with higher levels of prior knowledge may be better able to isolate question demands (e.g., causes of an illness), identify categories of texts presenting information meeting such demands (i.e., texts presenting causes), and focus on those sources from among a document set. Indeed, multiple source tasks have been identified as distinct

because of their reliance on analogical reasoning to generate inferences about texts and integrated information across sources. Thus, multiple source tasks may be more dependent on prior knowledge (Guthrie & Kirsch, 1987; Wenger & Payne, 1996), as students with higher levels of prior knowledge may have a greater store of information to draw on in identifying analogical structures within and across texts and in understanding texts in relation to other texts (Le Bigot & Rouet, 2007).

Prior knowledge supporting multiple text organization may be all the more important when students use multiple texts to respond to general (i.e., open-ended questions), as these offer less of a direct match between task demands and source content (Rouet, 2003). In open-ended question contexts, as in the present study, students may be more challenged in categorizing and linking texts to meet question demands—a process supported by their prior knowledge.

Differences in text processing for students high and low in prior knowledge have manifest in differences in time devoted to source use (e.g., Dreher & Guthrie, 1990; Rouet, 2003). For example, Le Bigot and Rouet (2007) examined students' total reading time, devoted to processing both the source menu and texts themselves. Reading times were significantly longer for low knowledge students, when reading both texts and menus, despite high knowledge students demonstrating better performance. Le Bigot and Rouet (2007) suggest that for students with high prior knowledge reading time is reduced in part due to understandings of the global structure of multiple texts (Hofman & van Oostendorp, 1999; McDonald & Stevenson, 1998; Rouet, 2003) or text organization. Students' with high levels of prior knowledge may also have greater experience with text

and document formats within a particular domain, reducing processing time (Rouet et al., 1997).

As in prior research, time on texts is used in the present study as a measure of text processing. Although earlier work has found students with higher knowledge to spend less time on text processing, this has largely been examined in the navigation of hypertext menus (e.g., Le Bigot & Rouet, 2007; Rouet, 2003). While high knowledge students may indeed spend less time on understanding text structure and organization, time on texts, in this study, is used as a metric of text processing and engagement during source access. Similarly, other studies have classified time on texts as a measure of cognitive effort expended (Bråten et al., 2014) corresponding to depth of processing.

Prior knowledge and source evaluation. Prior knowledge has been most extensively explored as impacting source evaluations and evaluative justifications across a variety of tasks and assessment contexts (Bråten et al., 2011; Rouet et al., 1996; Rouet et al., 1997). Indeed, starting with work comparing expert and non-expert samples (e.g., Wineburg, 1991), prior knowledge has been found to differentiate both the frequency and the nature of students' engagement in source evaluation. For instance, Rouet et al. (1997) in examining differences in the source evaluations of graduate history students, considered to be disciplinary experts, and graduate students in psychology, considered to be disciplinary novices, found a number of variations not only in the source evaluations history students generated but also in the types of justifications for source evaluation they offered. History graduate students drew on a wider range of justification criteria (i.e., those associated with content, source, and task-based considerations) than did psychology students, who focused on content related justifications, in explaining their reasoning for

source evaluation.

Despite the documented differences in the source evaluations of novices and experts, less is known about the effects of prior knowledge on the source evaluations of non-experts (i.e., the extent to which non-expert learners, differing in prior knowledge, show differential patterns or profiles of source evaluation). This is a promising area for examination given Bråten, Strømsø, and Salmerón's (2011) critique that studies examining prior knowledge have often contrasted expert and novice samples that demonstrate extreme differences in prior knowledge.

Bråten, Strømsø, and Salmerón (2011) examined the extent to which evaluations of source trustworthiness and the justifications for these evaluations differed for undergraduate students with varying levels of prior knowledge (i.e., non-experts). Limited differences in the trustworthiness ratings of students with low and high levels of prior knowledge were found (Bråten et al., 2011). In particular these two groups differed only in their evaluations of the least reliable source in the document set, a presentation from an oil company on the topic of climate change. In terms of the types of justification criteria cited, students in both the high knowledge and low knowledge groups most commonly reported determining source trustworthiness based on content. However, high-knowledge students seemed to distinguish between the various judgment criteria more so than did low-knowledge students. For the low knowledge students, there were no significant differences in justification criteria applied, beyond the frequency of judgments based on content. For high-knowledge students, trustworthiness ratings were based significantly more on publisher and source type than they were on author and date of publication.

Despite Bråten et al.'s (2011) somewhat limited findings about differences in the source evaluations of high and low knowledge students, there is good reason to investigate further. First, as Bråten, Strømsø, and Salmerón (2011) were interested in person-centered analyses, they created high and low prior knowledge groups via a median split, potentially limiting some of the variance in their sample's prior knowledge (Bråten et al., 2011). Further, the finding that high knowledge students better distinguished between justification criteria, particularly those that are document information related (i.e., publisher, source type) is significant.

There are two potential reasons why students with high levels of prior knowledge may be better able to engage in sourcing and therefore text evaluation as compared their low knowledge counterparts (Bråten et al., 2011). First, it may be the case that students with high prior knowledge are aided in their text comprehension and therefore have reduced cognitive load, allowing them to expend cognitive effort on evaluating the trustworthiness of sources. An alternate explanation may be that only those students high in prior knowledge have the requisite skills to comprehend texts in a way that allows them to judge the trustworthiness of the information within them. Further, learners with high levels of domain knowledge may have greater document familiarity or knowledge of domain-specific source types and source use practices; this knowledge may aid students in attending to document information and in developing source evaluations based on said information. Additional reasons why high and low knowledge students differ in their source evaluations may be that learners low in prior knowledge have difficulty distinguishing the criteria along which to base their source evaluations or are easily seduced by sources that seem trustworthy but indeed are not (Bråten et al., 2011).

Differences in the source evaluations of non-experts with varying levels of prior knowledge have been examined only to a limited extent, such as in Bråten, Strømsø, and Salmerón's (2011) exploratory study. However, given that prior knowledge has been theorized to be facilitative of source evaluation in a number of ways, it is assessed and examined in the present study as both associated with students' multiple source use behaviors and as predictive of task performance.

Stance on Target Topic

As compared to prior knowledge, attitudes, have been examined to a much more limited extent in the multiple source use literature. Attitudes may be defined as evaluative judgments about topics or objects (Crano & Prislin, 2006 in Van Strien, Brand-Gruwel, & Boshuizen, 2014, p. 101) and have been conceptualized as aligning with one of two sides of an issue (e.g., *pro* or *anti* capital punishment, Lord et al., 1979). As such, information presented through multiple texts has been classified as being either *belief consistent*, corresponding to students' pre-existing attitudes, or *belief inconsistent*, contrasting or conflicting with students' prior attitudes.

The belief consistency or inconsistency of information has been found to shape students' source selection, processing, and evaluation via mechanisms similar to those of prior knowledge. First, students have been found to be selective in choosing texts, preferring those sources they believe will confirm their prior attitudes (Van Strien et al., 2014).

Further, as is the case with prior knowledge, prior attitudes function as a schema in shaping the situation model learners construct (Wiley, 2005). As learners integrate new information into existing schema, resulting situation models have been found to be

biased toward belief consistency (Maier & Richter, 2013). This may be because reconciling and integrating conflicting points of view demands great cognitive load or effort or because students experience cognitive dissonance when information conflicts with prior beliefs. Schema theory further suggests that students' attitudes may function as a type of "knowledge structure," guiding the selection of information in texts, serving as a lens through which new information is interpreted, and supporting information integration (Pratkanis, 1989 in Maier & Richter, 2003).

Moreover, students have demonstrated biased evaluations of belief consistent arguments, considering them to be more compelling than parallel arguments that are belief inconsistent (Lord et al., 1979; Taber & Lodge, 2006); this differential evaluation has been termed *biased assimilation* (Kobayashi, 2010). Munro and Ditto (1997) have found students' positive evaluations of belief consistent arguments to include judgments of *quality* and *convincingness*.

In addition to students' preferential selection, processing, and evaluation of belief consistent information, researchers have identified a correspondent skepticism toward belief inconsistent information. First, in selecting sources, students may fail to attend to belief inconsistent information (Fischer & Greitemeyer, 2010; Fischer, Jonas, Frey, & Schulz-Hardt, 2005). For example, Brannon, Tagler, and Eagly (2007) asked students to rate the extent to which they wanted to read articles on "allowing abortion on demand." Strength of attitudes was associated with wanting to read attitude consistent information, with students being less interested in reading belief inconsistent information. Students being more limited in their access to belief inconsistent information has been termed *selective exposure* (Brannon et al., 2007). Further, when processing belief inconsistent

information, students have been found to be more critical and less accepting of arguments conflicting with their views (Lord et al., 1979). This may be because learners accept belief consistent information uncritically or because belief consistent argument evaluations serve to *anchor* the, then, negative evaluations of belief inconsistent arguments (Taber & Lodge, 2006).

Attitudes have also been associated with learners' task performance, typically on text-recall tasks. As with prior knowledge, students have been found to have better recall of belief consistent, as compared to belief inconsistent, information (Levine & Murphy, 1943; Wiley, 2005). Improved recall may stem from attitudes directing students' attention to particular information in texts or from belief consistent information being more readily integrated into existing schema (Maier & Richter, 2013). Improved belief consistent information recall has been adopted as evidence of the *congeniality hypothesis* (Eagly & Chaiken, 1993). The congeniality hypothesis postulates that students better recall information that, "supports, confirms, or reinforces their evaluations of social, political, and personal issues" (p. 64) as compared to conflicting information (Eagly, Chen, Chaiken, & Shaw-Barnes, 1999).

However, support for the congeniality hypothesis is not absolute. In a meta-review of studies examining the congeniality hypothesis, Eagly et al. (Eagly, Kulesa, Chen, & Chaiken, 2001) found its effect to be quite small. Additionally, it has been found that arguments that are incompatible with prior attitudes, although rated lower in quality, have also been subjected to longer scrutiny and have elicited more relevant thoughts from students, particularly toward their refutation (Edwards & Smith, 1996). This suggests the presence of a *disconfirmation bias*. Disconfirmation bias refers to

students' tendency to allocate more attention and cognitive effort to information that is belief inconsistent while quickly and uncritically accepting information that agrees with their point of view (Taber & Lodge, 2006).

Thus longstanding findings offer conflicting views with regards to the role of attitudes in text processing or at least text recall. This may stem from a number of limitations in the attitudes research. First, investigations of attitudes' effects on source use have most commonly examined students' performance on text recall or memory tasks. To the extent that students have been found to demonstrate greater cognitive engagement with belief inconsistent information (Eagly et al., 2001; Taber & Lodge, 2006), recall tasks may be insufficient to reflect this depth of processing. Rather, task-directed multiple source use, whereby students are required to consult multiple texts in order to formulate a position on a given issue, may better capture learners' reasoning about texts.

Further, in assessing attitudes, the issues presented have been bipolar (i.e., pro or con, Eagly et al., 1999) and typically, students have been classified dichotomously as supporting one position or another. As a result, the extent to which individuals' beliefs align with a particular position has been conflated with extremity of views (Judd & Kulik, 1980). In other words, there has not been a distinction drawn between attitude content (e.g., pro death penalty) and attitude strength. Generally, recall and text processing favor extreme rather than more moderate attitudes, suggesting the need to look at attitudinal strength or valence rather than absolute belief. Indeed, many of the findings regarding attitudes and information processing are considered to result from motivated processing toward attitude defense (Taber & Lodge, 2006). This is predicated

on attitudinal strength (Eagly et al., 1999; Eagly & Chaiken, 1993; Krosnick & Petty, 1995) or investment in one's beliefs.

As an example, Van Strien, Brand-Gruwel, and Boshuizen (2012) separated students into two groups, with one group holding a skeptical attitude toward climate change and the other group exhibiting a neutral attitude toward the issue. However, these groups were not balanced with regard to attitudinal position. While students who were classified as climate change skeptics may have held strong beliefs on the issue, students in the neutral group did not represent an opposing position; rather these were students not expressing strong beliefs. In the present study, belief valence rather than specific belief content is examined as potentially associated with multiple source use.

It may be particularly important to investigate students' attitudes within the context of information use on the Internet. In digital contexts, where students are freely able to select among a multitude of texts (Shapiro & Niederhauser, 2004), there is an increased probability that students will be able to easily access belief consistent information and navigate away from attitude inconsistent information while still having a plethora of sources to choose from. As a result, students may spend significantly more time processing information consistent with their prior attitudes (Graf & Aday, 2008; Knobloch-Westerwick & Meng, 2011) rather than being challenged by conflicting viewpoints.

In this study, rather than examining attitudes, which assume that students hold a well-formed and consistent system of beliefs accompanied by schematic knowledge (Wood et al. 1995), students' *stance* on the target issue was considered. Chaiken (1980) distinguishes between two types of attitude-based information processing *systematic* and

heuristic. Systemic processing occurs when learners deliberately consider and deeply process persuasive arguments in light of their existing attitudes. Heuristic processing occurs when students exert comparatively limited cognitive effort on judging message content and rely on surface cues to evaluate information. While systematic processing is comprehensive and detail-focused, heuristic processing uses simple rules to make general judgments when processing information (Chaiken, 1980). Given the naiveté of students' knowledge and attitudes regarding the Arab Spring in Egypt, examining learners' stance may best be described as considering their potential heuristic judgments of information based on general views on the topic.

Much of the research on attitudes in text processing has focused on systematic belief change, whereby readers may deliberate over persuasive information in light of existing attitudes, and, after reflection, may experience a change in attitudes (Buehl, Alexander, Murphy, & Sperl, 2001; Chaiken, 1980). At the same time, work on MSU suggests that when evaluating sources, learners often form heuristic judgments and, in part, may be forced to do so due to the cognitive complexity presented by multiple texts (Brem et al., 2001; Gerjets et al., 2011). This presents a seeming conflict between the attitude-driven, deliberative evaluation expected of students and learners' superficial judgments of information. In light of this conflict, particularly in a sample expected to be novice with regard to the topic, rather than examining attitudes it seemed more appropriate to examine stance, or students' general positioning with regard to the target prompt.

Chaiken (1980) suggests that heuristic based processing takes hold when students experience economic concerns regarding the amount of cognitive effort necessary to

process texts in attitude-referent ways and to complete a task. Multiple source tasks may push students' economic concerns to the forefront. Specifically, using multiple texts to understand a complex issue about which learners have limited knowledge may require a great deal of cognitive effort. As a result, students may be stingier in expending effort on systematically scrutinizing information in light of their attitudes, particularly when these are not well developed. More generally, multiple text tasks have been defined as being task driven (Rouet & Britt, 2011), as such, students may be more focused on processing sources in task-focused rather than attitude-referential ways. Thus, when presenting learners with a complex multiple text task on a topic about which students had limited knowledge, assessing topic stance, rather than attitudes, may be sufficient to examine heuristic-based text processing.

To the extent that both prior knowledge and stance may lead to selective text processing and evaluation, an additional individual difference factor was considered, students' self-reported domain general source evaluation behaviors.

Domain General Source Evaluation Behaviors

Contrary to the well-established role of domain knowledge in multiple source use (e.g., Britt & Aglinskis, 2002, Rouet et al., 1996; Rouet et al., 1997; Wineburg, 1991), students' *experiences* with multiple source use have not been fully considered. At the same time, there is initial support for examining students' multiple source use habits, in general, and source evaluation behaviors, in particular.

A wide variety of self-report measures have been used to examine students' behaviors associated with multiple source use. While some work has examined students' multiple source use experiences in a general manner, targeting a variety of skills and

competencies associate with multiple source use (Kunkel, Weaver, & Cook, 1996; Mason et al., 2010b; Walraven et al., 2009), the majority of studies examining self-reported multiple source use behaviors have asked students to report their general practices with regard to source evaluation. For instance, Burton and Chadwick (2000) surveyed college students to identify the criteria they applied in evaluating online sources and to identify the type of instruction they had received regarding source evaluation. In reporting the criteria important in evaluating both Internet and library sources used for writing research papers, students prioritized source availability, ease of location and understanding, and up-to-datedness. At the same time, students rated criteria related to author credentials, publication reputation, and peer-review much lower (Burton & Chadwick, 2000). Further, Burton and Chadwick (2000) found learners to report applying consistent source evaluation criteria across various types of academic assignments.

Head and Eisenberg (2010), likewise using a survey methodology, identified the criteria students reported using for web content evaluation. Most commonly, students considered source currency, author's credentials, the URL or website domain, and interface design. Based on post-hoc interviews, Head and Eisenberg (2010) concluded that for most students source evaluation did not constitute a dedicated step in the multiple source use process, rather students evaluated sources only to determine which information to include in written assignments. Of course, such surveys provide limited information about the actual amount of time and effort students devote to source evaluation. Although students' domain general source evaluation behaviors have been found to be an important descriptive characteristic of undergraduate populations, limited research has directly associated these reported behaviors with learners' performance on

multiple source tasks. In the present study, domain general source evaluation behaviors are examined as associated with source evaluation and performance on a multiple text task.

While the importance of each of these three learner characteristics (i.e., prior knowledge, stance, and domain general source evaluation behaviors) for multiple source use has been considered in prior research, all three of these individual difference factors have yet to be analyzed within a single model. Nonetheless a limited number of studies have sought to adopt a more comprehensive approach to modeling multiple source use and task performance.

Empirical Examinations of Models of Multiple Source Use

The most comprehensive set of examinations of models of multiple source use have been carried out within the framework of the Information Problem Solving on the Internet (IPS-I; Band-Gruwel, Wopereis, & Walraven, 2009) model of multiple source use. In parallel with the MD-TRACE model, the IPS model identifies core source use processes students perform when locating specific information to respond to discreet queries. This model examines students' abilities to perform three constituent skills: *searching for information*, *scanning information*, and *processing information* (Brand-Gruwel, Wopereis, & Vermetten, 2005). Although the majority of studies undertaken as part of the IPS framework have focused on examining each of these three constituent skills individually, some studies have considered the multiple source use process more holistically.

For example, Walraven et al. (2009) investigated the full ISP model to determine the amount of time students devoted to constituent skills (i.e., searching for, scanning, and processing information) and the frequency with which they engaged in these

behaviors. Across discrete questions in various domains, students devoted the most time to searching for information with the least time spent on processing information. This pattern also reflected the frequency with which these skills were invoked: students most often searched for information, then scanned information, and used skills associated with processing information with considerable less frequency. While Walraven et al. (2009) examined and compared each of these constituent skills to one-another, these were not jointly examined nor were they associated with students' performance. However, a regression using select source use indicators (i.e., number of sources accessed and evaluation criteria cited) as predictors of performance in responding to discrete questions was found to be non-significant. Although the regression model examined did not include individual difference factors, students' WWW Knowledge (i.e., academic experience) was assessed via a questionnaire and presented descriptively.

More recently, Bråten, Anmarkrud, Brandmo, and Strømsø (2014) introduced a model of individual difference factors, multiple source use behaviors, and task performance. Citing the lack of a similarly comprehensive model, Bråten et al. (2014) examined both direct and indirect effects of learner characteristics and source use behaviors on task performance. Individual difference factors considered were prior knowledge, individual interest, need for cognition, and an epistemic belief orientation endorsing a justification for knowledge by multiple sources. Multiple source use processing factors examined were effort (i.e., time students devoted to source use), students' situational interest, and self-reported use of deeper level strategies indicative of cross-text elaboration (i.e., the comparison, contrast, and integration of information across multiple texts).

Overall, Bråten et al. (2014) determined that while processing variables (i.e., effort, deep processing strategies) had a direct effect on multiple text comprehension, individual difference factors (e.g., prior knowledge, need for cognition) had only an indirect effect and were mediated by processing variables. In fact, prior knowledge was the only individual difference factor that had a direct effect on multiple text comprehension. While these models offer promising insights into the inter-relations among individual difference factors, multiple source use processing, and task performance, more work is needed to better align students' demonstrated source use with theoretical conceptions of this process. Such a task is the focus of the present research.

CHAPTER 3: METHODOLOGY

This study had five primary goals. First, I was interested in describing the nature of students' multiple source use when responding to an open-ended task. Second, this study examined the role of multiple source use behaviors in predicting response quality. Next, it considered the relation between individual difference factors and multiple source use behaviors as well as the relation between individual difference factors and task performance. Finally, a full model of multiple source use was examined, whereby individual difference factors and multiple source use behaviors were used as predictors of open-ended response quality.

This chapter begins with a discussion of participants as well as efforts undertaken to ensure adequate sample size. Next, this chapter describes measures used to assess individual difference factors, capture students' multiple source use behaviors, and evaluate response quality.

Participants

This study included data from 197 undergraduate students from a large mid-Atlantic university. The sample was 65.48% female ($n=129$) and 29.95% male ($n=59$). Participants were on average, 20.47 years old ($SD=2.08$). The sample was 49.23% White ($n=97$), 19.29% Asian ($n=38$), 16.24% African American ($n=32$), and 3.55% Latino ($n=7$). Two students reported their race/ethnicity as "Other" and 6.09% of students ($n=12$) self-reported as either biracial or mixed race.

The sample included a range of class standings: 21.32% of students were freshman ($n=42$), 21.83% were sophomores ($n=43$), 27.41% were juniors ($n=54$), and 19.80% were seniors ($n=39$). Additionally, 5.08% of students ($n=10$) were taking

courses beyond their senior year. Participants represented a variety of majors. Most students, 56.06%, were majoring in the social sciences (n=111), 27.27% of students were majoring in the natural sciences (n=54), 5.05% of students were humanities majors (n=10), and 6.60% of students had not declared a major (n=13). Participants reported an average GPA of 3.26 ($SD=0.53$). Nine students did not report demographic information.

Undergraduate students were concurrently recruited for participation using three methods. First, students were recruited from courses where instructors were offering extra credit for participation. Second, participants were recruited using the SONA Psychology Research system, wherein students in select psychology courses are required to participate in research studies for course credit. Third, the paid SONA system was used. Participants signing up through the paid SONA system received \$10 for participation. As such, the participants in this study were a convenience sample.

Sample Representativeness

The current sample was compared to the overall undergraduate student population enrolled in Fall 2013 at the university where the study took place. Table 1 presents a comparison of students in my sample to those enrolled at university at large.

Chi-squared one variable tests were used to examine whether the recruited sample was proportionate to the overall population with regard to gender, race/ethnicity, and class standing. Females were significantly overrepresented relative to males, $X^2(1) = 36.62, p < 0.001$. Race/ethnicity was dichotomized as minority (i.e., African American, Asian, Hispanic/Latino, Mixed Race) or not (i.e., White). Racial categories were collapsed in accordance with institutional reporting and to prevent violations of chi-squared assumptions (i.e., 5 cases per cell). Minority students were overrepresented

relative to White students, $X^2(1)=4.17, p<0.05$. Based on an examination of residuals, seniors were underrepresented relative to freshman and special students were overrepresented, $X^2(4) =10.54, p<0.05$.

Table 1

Demographic Comparison of Sample to Overall University Population

		University (Fall 2013)	Sample
Gender	Male	53.4% (n=14223)	29.95% (n=59)
	Female	46.6% (n=12435)	65.48% (n=129)
Average Age		21	20.47 <i>SD</i> =2.08
Race/Ethnicity	White	53.4% (n=14226)	49.24% (n=97)
	Black/African American	12.5% (n=3,326)	16.24% (n=32)
	Asian	15.4% (n=4,117)	19.29% (n=38)
	Hispanic	8.8% (n=2335)	3.55% (n=7)
	American Indian/ Pacific Islander	0.2% (n=54)	0%
	Two or More	3.6% (n=956)	6.09% (n=12)
	Unknown	2.9% (n=785)	4.57% (n=9)
	Minority Status	40.5%	46.19%
Class Standing	Freshman	16.8% (n=4468)	21.32% (n=42)
	Sophomore	22.3% (n=5926)	21.83% (n=43)
	Junior	27.1% (n=7180)	27.41% (n=54)
	Senior	30.0% (n=7966)	19.80% (n=39)
	Post-Bac/Special	3.6% (n=947)	5.08% (n=10)
	Undergraduate		

While demographic factors were not of interest in this study and were not included as part of the theoretical model guiding present work, these imbalances present potential challenges to the generalizability of findings. Demographic factors were statistically controlled for, with supplemental models presented in Appendix A, but not included in focal analyses specific to each research question, as this was considered to be beyond the scope of the present study. Limits to the generalizability of findings are discussed in Chapter 5.

The extent to which participants' areas of study were representative of the overall university population could not be ascertained, as descriptive information regarding students' majors are not reported by the university.

Sample Size Estimation

The target sample size for the present study was determined by conducting a priori power analyses (i.e., looking at the degrees of freedom required for various planned analyses to detect effects of various sizes, Cohen, 1992). Target sample size was computed using the software G*Power 3.1.9 (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007). The most taxing analyses in this study, in terms of sample size, were multiple regression models. The full regression models examined included 9 predictors (i.e., learner characteristics: prior knowledge attitudes, domain general source evaluation behaviors; and multiple source use behaviors: total number of sources access, total time on texts, accessing document information, discrimination in trustworthiness ratings, average source usefulness ratings, average source interestingness ratings).

In specifying desired effects size, prior literature was examined. In the model most closely resembling my own, learner characteristics and source use behaviors were found to explain 40% of variance in students' task performance, described as a large effect (Bråten et al., 2014). However, the individual effects of predictors in that study were classified as small to medium. As I was primarily concerned with overall model fit, a large effect size was expected. Yet, power analyses were run conservatively, based on a medium effect size, to ensure adequate sample size. To my knowledge, no other work has been comparable to the present study in examining a full model of multiple source

use, including both individual difference factors and multiple source use behaviors as predictors of response quality.

A multiple regression with $\alpha = 0.05$ and $\pi = 0.80$, able to detect a medium effect size (i.e., $f^2 = 0.15$), as recommended by Cohen (1992), would require a sample of 114 students. To account for possible attrition (conservatively estimated at 20%), the target sample for this study was set at 137 participants.

This sample size estimation was consistent with a number of heuristics that have been proposed in prior research. For example Stevens (1996) recommends including 15 subjects per predictor for social science research, which would have resulted in a target sample size of 135. Tabachnick and Fidell (2001) recommend including $50 + 8k$ participants, where k corresponds to the number of predictors in the model. With nine predictors in the full model, this necessitated a sample size of 122 participants. The number of participants in the recruited sample was deemed sufficient to meet these sample size demands.

A challenge in estimating necessary sample size was that students in this study were not required to access all of the sources in the library. As a result, participants did not necessarily have evaluations corresponding to each source. Those students not accessing a particular source appeared to have missing data and would have been removed from the regression model via listwise deletion. This limitation was mitigated in two ways. First, during data collection, participants' source access was monitored to ensure that all of the sources were accessed by a sufficient number of students. Second, rather than using source specific metrics as predictors (e.g., trustworthiness ratings of a newspaper article), overall indices of source use (i.e., discrimination in trustworthiness

ratings, average usefulness rating, average interestingness rating), based on those texts that students did access, were computed.

Undergraduate Students as a Target Sample

This study examined multiple source use in a sample of undergraduate students. Undergraduates were the population of interest for a variety of reasons related to students' developmental stage, the academic expectations placed on college students, and the context of today's learners' classroom experiences. First, college has been considered to be a formative academic experience for students and therefore has been targeted as a period during which to examine learning, generally (Kuhn, 1991; Perry, 1970, 1981), and multiple source use, more specifically (Bråten et al., 2009; Bråten et al., 2011; Rouet et al., 1996). In college, students may be more commonly exposed to controversial topics that require them to engage with and evaluate sources with conflicting points-of-view (Rouet et al., 1996). Undergraduates also more commonly engage with a wide variety of academic sources (e.g., scholarly books, journal articles) as compared to their high-school counterparts, who may rely primarily on textbooks (Rouet et al., 1996). Engaging with different types of sources may require students to exhibit greater facility in selecting, processing, and evaluating this variety of texts (Britt et al., 1999).

In general, in college, students are regarded as developmentally mature thinkers capable of the higher-order cognitive processes associated with multiple source use, including the weighing of competing claims and the critical examination of evidence (Hofer & Pintrich, 2004; King & Kitchener, 2004; Kitchener & King, 1981). At the same time, recent investigations of college students' academic behaviors, particularly when conducting research on the Internet, have found that students are often not very skilled in

multiple source use, particularly with regard to source evaluation (Metzger et al., 2003). For example, Grimes and Boening (2001), examining undergraduates' essays, determined that students evaluated web resources only superficially, while Graham and Metaxas (2003) reported that students seldom corroborated information found on websites with additional sources.

These limitations in multiple source use manifest at the same time that university curricula have placed an increased emphasis on research skills (Metzger et al., 2003). Indeed, students' demonstrated multiple source use skills have been found to fall short of faculty expectations (Grimes & Boening, 2001). College constitutes a critical period in which to examine students' difficulties with multiple source use not only because this is a time when such deficiencies manifest but also because students' abilities to engage with multiple sources in complex ways may have real-world consequences (Kammerer, Bråten, Gerjets, & Strømsø, 2013; Kim & Millis, 2006). For example, college-aged learners may be called upon to engage in multiple source use to make choices based on competing evidence when making political decisions (e.g., Kushin & Yamamoto, 2010) and researching personal health information (e.g., Kienhues et al., 2011).

Finally, examining multiple source use in an undergraduate sample served as a valuable contribution to the research literature. Bråten, Strømsø, and Salmerón (2011) highlight the examination of non-expert (e.g., undergraduate students) source use as an important direction for work on multiple source use and a departure from expert-novice studies, extensively considered in prior research (e.g., Rouet et al., 1997; Wineburg, 1991).

Non-Expert Sample Status

Two facets of expertise may be considered in describing the current sample. Specifically, students' expertise with regard to the topic as well as with regard to multiple source use may be examined. Ericsson and Smith (1991) define *experts* as, "outstanding individuals in a domain" (p. 3), while Alexander (2003) has defined experts as individuals with broad and deep knowledge of a domain who also contribute to the development of new knowledge. It is reasonable to assume that the undergraduate students in my sample were non-experts. However, non-experts are definitionally an expansive and varied category of individuals, including students acclimating to a domain, with limited and fragmented knowledge, and learners demonstrating competency in a given field, with a schematic stores of knowledge (Alexander, 2003). Undergraduate students, who are non-experts, have been distinguished according to their prior knowledge. Bråten et al. (2011) administered a 17-item prior topic knowledge assessment to learners (i.e., with a maximum score of 17). Students scoring a mean of 5.52 points out of 17 were classified as *low knowledge* and those students scoring a mean of 9.00 points were classified as *high knowledge*, according to a median split. The students in the current sample had an average prior knowledge score of 2.32 ($SD=2.62$) on a 7-point scale. Using Bråten et al.'s (2011) classification guidelines, these students may best be classified as a low knowledge sample.

To my knowledge, no work has defined what may constitute expertise with regard to multiple source use. However, Rouet et al.'s (1997) comparison of experts (i.e., graduate students) across domains, suggests that multiple source use expertise has a strong domain-specific component. While the students in this sample reported engaging

in domain general source evaluation behaviors with moderate frequency ($M=4.45$, $SD=1.22$ on a 7-point scale), particularly given their limited prior knowledge, students were likely not expert in topic-specific source use.

Measures

The study was carried out over two data collection sessions. In Session 1, demographics and learner characteristics (i.e., prior knowledge, stance, source evaluation behaviors) were collected. In Session 2, participants completed a multiple source use task on the topic of the Arab Spring in Egypt. Participants completed Session 1, at a time and location of their choosing, while Session 2 was completed in a computer laboratory. Session 1 and Session 2 were separated due to concerns regarding the cognitive and time demands associated with study completion. All data were collected online via research study websites.

Session 1: Learner Characteristics

Prior to engaging in the multiple source use task, students were asked to report four metrics: (a) demographics, (b) prior knowledge, (c) topic stance, and (d) frequency of engaging in domain general source evaluation behaviors, assessed as part of an academic behaviors questionnaire. A part of the academic behaviors questionnaire, the *Credibility Assessment Scale*, was of particular interest in this study as a measure of domain general source evaluation behaviors.

Demographics. Students were asked to report demographic information to aid in describing the characteristics of the sample. Demographics were not included in primary analyses of students' multiple source use, although they are included as control variables in models presented in Appendix A. Participants were asked for commonly reported data

about their academic and non-academic backgrounds. Questions addressing students' academic backgrounds asked learners for their major, GPA, and class standing (e.g., sophomore). Non-academic demographic questions asked participants to report their gender, age, and race/ethnicity. The demographics questionnaire is included in Appendix B.

Prior knowledge. A seven-item identification measure was administered to assess students' prior topic-level knowledge of the Arab Spring in Egypt, the topic of the MSU task. Specifically, students were specifically asked to, "Please tell me about," pertinent people (i.e., Mohamed Morsi, Hosni Mubarak, and General el-Sisi.), places (i.e., Tahrir Square), and terms (i.e., Arab Spring, Muslim Brotherhood, and Tamarod) associated with the uprisings in Egypt, drawn from news stories and reports of unfolding events. The prior knowledge items were chosen to reflect the key terms students would encounter in reading and interacting with the texts included in the MSU task. Definitions were coded on a binary scale as correct or incorrect, for a maximum prior knowledge score of seven. Wineburg (1991) similarly used an *identification of terms* task as a measure of prior knowledge. The seven-item prior knowledge measure had a Cronbach's alpha of 0.91. Students' mean prior knowledge scores were 2.32 ($SD=2.62$) on a 7-point scale. The prior knowledge measure is included as Appendix C.

Stance. Students' stance with regard to the Arab Spring in Egypt was assessed via two items. Students were asked to report both their stance with regard to who should be in power in Egypt (i.e., Morsi or el-Sisi) and how strongly they identified with such a position. The first asked participants to report, "How strongly do you feel about events associated with the Arab Spring in Egypt" on a 7-point Likert scale ranging from, "I have

no opinion on the issue” to “I have very strong opinions on the issue.” The second item asked participants, “Who should the US support to hold power in Egypt.” Students were able to select from three answer choices: “Mohamed Morsi and the Muslim Brotherhood,” “General el-Sisi and the military” or “I don’t know.” Further an “Other” option was provided, wherein students were able to specify an alternate choice. This second question paralleled the MSU task students were asked to complete, providing a justification for why the US should support Mohamed Morsi, General el-Sisi, or another alternative. The first question was used in analyses as a metric of students’ attitudinal valence, while the second question was used descriptively to identify students’ initial stances on the task topic. The two items assessing stance are included as Appendix D.

A key limitation in this study is that stance was assessed via a single item rather than a scale. The use of single item scales is generally not recommended as it presents limitations with regards to reliability (Churchill, 1979). Additionally, assessing a construct using only a single item introduces added concerns regarding order of item presentation in relation to other measures administered and item wording (Dolbier, Webster, McCalister, Mallon, & Steinhardt, 2004). For instance ambiguity in phrasing or a disconnect between intended item interpretation and students’ conceptions of a particular questions may present serious validity issues. This limitation was assumed for a number of reasons. In general, the literature on multiple source use has assessed attitudes only to a limited extent and no existing attitude measures associated with the topic were identified. Those attitude measures found in literature were associated with attitude content (i.e., ascertaining a position as pro or con, such as pro death penalty) rather than attitude valence or strength of belief (e.g., Kobayashi, 2010; Van Strien et al.,

2012), as was the focus in the present study. Further, in prior work, when students have been presented with attitude scales these have been preceded by a paragraph providing background information explaining the target issue before assessing attitudes (e.g., Kobayashi, 2010; Sparks, 2013). This methodology was avoided in the present study, as I wanted all information about the Arab Spring to be selected by students from sources in the library.

The single item measure of stance was therefore piloted and time constraints associated with selecting a contemporary topic prevented piloting a full attitude scale (i.e., data had to be collected prior to a planned election in Egypt scheduled for May 26-28, which would have altered task-relevant information). Further, single item attitude assessments have been used in prior research (e.g., attitude extremity, Edwards & Smith, 1996) and may be appropriate when items explicitly ask for reports of a single, concrete attribute (e.g., interest, Bergkvist & Rossiter, 2007). In this case, I was not interested in multidimensional attitude components, rather just a single metric of student stance. In addition, single items may be appropriate to use when they are unambiguous to participants (Wanous, Reichers, & Hudy, 1997). In this case, the item was explicit in assessing general position on the topic. Single item measures have been used in prior multiple source research to ascertain students' initial attitudinal positions in relation to a given task (e.g., Sparks, 2013).

Domain general source evaluation behaviors. To capture students' domain general source evaluation behaviors, a nine-item *Credibility Assessment Scale* was administered to students as a part of a longer academic behaviors questionnaire. The credibility assessment scale asked students to report the frequency with which they

engaged in nine evaluative behaviors, “when looking for information for schoolwork.” These behaviors included, “check[ing] to see if the information is current” and “seek[ing] out other sources to validate the information” (Hargittai et al., 2010, p. 477). Items were presented in random order to participants. Students were asked to report the frequency with which they engaged in each evaluative behavior using a scale ranging from one to seven, with one corresponding to “never” and seven corresponding to “very often.”

The credibility assessment questionnaire was based on Hargittai et al.’s (2010) eight-item measure. The Hargittai et al. (2010) credibility assessment scale was selected as the preferred measure for assessing students’ domain-general source evaluation behaviors as it has been used in large-scale surveys of diverse undergraduate populations and validated with observations of students’ source use. Further, the scale was brief and included items corresponding to source evaluation behaviors about which data could be collected during task completion (e.g., “check to see who the author is”). A similar credibility assessment scale, having one additional item, was found to have strong reliability ($\alpha = 0.92$, Metzger et al., 2003). Sample-specific scale reliability was established at Cronbach’s $\alpha=0.86$ for nine items. The credibility assessment scale is included as Appendix E.

The *credibility assessment* scale was nested in a more general survey of students’ academic behaviors to mask the purpose of the study. The more general assessment of students’ academic behaviors consisted of three additional scales: (a) seven items capturing students’ engagement in Internet-based learning activities (Bråten & Strømsø, 2006); (b) eight items asking participants to report the frequency with which they use various information sources (e.g., books, websites) when completing schoolwork

(Metzger et al., 2003); and (c) eight items asking students to rate the credibility of each of these information sources (Metzger et al., 2003). The full behavioral assessment scale is included as Appendix F.

Session 2: Multiple Source Use

While participants were asked to report learner characteristics prior to task engagement, all subsequent measures of multiple source use process and product were assessed as students completed a multiple source task within an online interface. Session 2 was completed by students online, in a computer laboratory at the university, with the researcher present. First, the nature of the multiple source use task is described. Next, the specific assessments that were used to capture students' multiple source use are detailed.

Multiple source use task. The multiple source use task that participants were asked to complete had two key parts: the research phase and the response phase. During the *research phase*, students were provided with a prompt initiating source use and a library of six texts, varying in source type. Students received the following instructions: "Please answer as you would if assigned to write a brief essay in response to this prompt for an academic class. In responding to the prompt you will be asked to take a position (i.e., in support of Mohamed Morsi or General el-Sisi or an alternative) as well as to provide specific evidence to support your position." Further students were directed to take notes while they researched and told that they would not be able to access any sources in the library while composing their responses. Participants were able to use their notes in writing their answers. Finally, students were told that there was no time limit for task completion. Appendix G includes a screenshot of directions provided as a

part of the study interface. Once students determined that they had sufficient information necessary to reply to the prompt, completing the research phase, they moved to the *response phase*. The response phase asked students to first, position themselves with regard to the prompt (i.e., to designate whether they were in favor of Morsi, el-Sisi, or an Other option), and then, to provide specific arguments and evidence in support of their position. Appendix H displays a screenshot of the response page.

Following the response phase, students were asked to complete an evaluation task, ranking each of the sources they accessed and justifying these ranking. This evaluation task was not examined as part of the current study as the focus in this research was on students' behaviors during source use and resulting task performance. Finally, students completed a post-task engagement measure, the results of which are presented descriptively in Chapter 4.

The research and response phases were separated for two primary reasons. First, in much of the work on multiple source use students have been asked to compose responses without referring to the source library (e.g., Bråten et al., 2011; Britt & Aglinskias, 2002). In part, this may be due to researchers' interest in capturing students' memory for sources when composing responses (e.g., Stadter & Bromme, 2007, 2008). Second, separating the research and response phases was intended to encourage students to engage in the source use process to a greater and more deliberative extent. The aim was to draw students away from satisficing considerations of what would be a "just satisfactory" response or where the "right" answer may be located in text; goals that have generally been found to be pronounced in students' source use (Wallace, Kupperman, Krajcik, & Soloway, 2000). Rather, asking students to engage in the research process in

preparation for, but independent of, response composition, was intended to better facilitate learners' determinations that they had an *information need*, the motivation for accessing sources, and decisions that their information needs had been met. More singular attention to information need may have resulted in more reflective or deliberate source engagement.

Task topic. The prompt initiating students' multiple source use was: "Should the United States support General el-Sisi and the military regime or Mohamed Morsi and the Muslim Brotherhood?" While a great many investigations of students' multiple source use have been nested in the domain of history (Wiley et al., 2009), referencing events in Classical and Modern history, the topic selected for the present study (i.e., the Arab Spring in Egypt) constitutes a contemporary political issue. Across time periods, students have been asked to examine topics as historical *controversies* or to engage with disagreeing or conflicting interpretations of events in history (e.g., Perfetti, Britt, Rouet, Georgi, & Mason, 1994; Rouet et al., 1997). Topics may be considered controversial when events may be understood from a variety of potentially conflicting perspective (Britt & Aglinskis, 2002) or when tasks prompt students to consider differing interpretations of events. The task selected for this study asked students to grapple with a *controversy*, as there is significant disagreement about how events in Egypt should be interpreted and which course of action the United States should pursue (e.g., Goodwin, 2011; Snider & Faris, 2011). Specifically, the task included a prompt asking students to align themselves with one of two opposing views on the issue based on texts presenting a variety of perspectives on the Arab Spring in Egypt.

Typical multiple source use studies addressing history topics have asked students

to draw on a variety of source types, both primary and secondary, written by both laymen and experts (Rouet et al., 1996). Likewise, in this study, multiple source types, both primary and secondary texts, from a variety of authors were provided to support students' reasoning about a contemporary event. Despite the similarities between the chosen contemporary political topic and more traditional history topics, some differences did emerge, particularly in the nature of the sources that could have been presented to students as a part of each task, as evident in Table 2.

Selecting a topic in contemporary politics allowed for the inclusion of more contemporary source types in the digital library. In particular, using an ongoing political topic allowed for the inclusion of digital and Web-based sources (i.e., sources created in digital contexts for online consumption, like blogs or Wikipedia) in the source library, as well as traditional print sources delivered electronically (e.g., newspaper). Web-based sources require further examination in the research literature as they are both evocative of their print-based predecessors (e.g., encyclopedias) and include new digital properties (e.g., hyperlinks, Coiro, 2003b). The topic of the Arab Spring in Egypt was specifically selected because it constitutes an event in which digital sources (e.g., Twitter) have been considered to play an especially important role (Khondker, 2011; Lotan, Graeff, Ananny, Gaffney, & Pearce, 2011).

Moreover, selecting a topic addressing an ongoing political issue allowed for the opportunity to confront students with a complex real-world issue of current importance. It was expected that students would have some familiarity with the topic of the task (i.e., the Arab Spring in Egypt) as it had been featured prominently in the news and public discourse since 2011. At the same time, it was thought that students would not

necessarily have strong attitudes or beliefs about the topic, as it concerns international, rather than domestic, events in which the United States has no direct involvement. Selecting a topic about which students likely did not hold strong beliefs and in which learners likely had limited personal involvement (e.g., unlike the wars in Afghanistan or Iraq) was intended to encourage students to respond to the prompt based on an evaluation of information presented in sources rather than based on their feelings or opinions. As expected, students did not report holding strong opinions on the issue ($M=2.05$, $SD=1.54$, on a 7-point scale).

Table 2

Comparison of Traditional History Topics Used in Prior Research and the Contemporary Political Topic in the Current Investigation

<p style="text-align: center;">Traditional History Topics Prior Studies</p> <p>Gulf of Tonkin incident and Vietnam War (Britt & Aglinskas, 2002); Construction of the Panama Canal (Rouet et al., 1996); Revolutionary War Battle of Lexington (VanSledright, 2002; Wineburg, 1991) Fall of the Roman Empire (Wolfe & Goldman, 2005)</p>	<p style="text-align: center;">Contemporary Political Topic Current Study</p> <p>Should the United States support General el-Sisi and the military regime or Mohamed Morsi and the Muslim Brotherhood?</p>
<p>Similarities</p> <ul style="list-style-type: none"> • Present multiple sources and source types, written from multiple viewpoints • Include both <i>primary</i> and <i>secondary</i> sources • Include sources authored by both experts and non-experts • <i>Controversial</i> – present differing explanations or interpretations of events • Explain events in terms of both individual actions and broader societal, economic, political, or cultural trends 	
<p>Differences</p>	
Events transpired in the past	Events ongoing
Sources may be produced during a longer span of time (from the time of event to contemporary historical writing)	Both participant accounts and expert accounts produced contemporarily, prior to event resolution
Texts are traditional print-based materials	Texts may be traditional print-based materials or digital sources, created digitally for online consumption (e.g., Twitter, blogs)
Students perceive events as resolved and know a popular explanation of events, even though there may be expert disagreement	Situation is unfolding and outcome unknown; less commonly accepted story to be told
Events are resolved; secondary sources may be written based on known outcomes and hindsight	Events still unfolding; both primary and secondary sources are writing without known resolution, contemporaneously with events
Primary source material may be limited and may be selected and verified by historians	Great deal of un-culled primary source material
Tertiary sources (e.g., textbooks) common	Tertiary sources may not be available

Task formulation. A task asking participants to use multiple texts in formulating arguments to support or oppose a controversial prompt (i.e., a prompt proposing

disagreeing or divergent interpretations of events) is typical of tasks used in prior research (e.g., Britt & Aglinskas, 2002). Indeed, studies of multiple source use have asked students to compose summaries, narratives, arguments, explanations, and opinion essays based on multiple texts (e.g., Gil et al., 2010a, 2010b; Wiley & Voss, 1999). Across a variety of task conditions, composing an argument-based essay, as compared to a summary or narrative, has been found to result in deeper source engagement and better performance on assessments of text comprehension and essay quality (e.g., Wiley & Voss, 1999). In the current task students were asked to produce arguments in favor of or in opposition to the target prompt.

Texts. Six sources were provided in the digital library, representing a variety of source types. A great variety of source types have been used in investigations of students' multiple source use both in the domain of history and beyond. These sources have included participant accounts of events (e.g., diaries), expert interpretations (i.e., historian's essays), and period-specific official documents (e.g., treaties; Britt & Aglinskas, 2002; Rouet et al., 1996; Wineburg, 1991). This variation in source types has reflected differences in author expertise and intentionality (i.e., source purpose), as well as in period of production or publication. Included sources have been written contemporarily or long after events have occurred (Rouet et al., 1996; Wineburg, 1991). The selection of a contemporary historical topic meant that while source type, author expertise, and intentionality varied across texts, all sources were written within the same general time period.

Strømsø, Bråten, and Britt (2011) outline four standards they applied when selecting sources, guidelines that were adopted in this study. First, sources were selected

to address a topic about which learners would have some, but not extensive, prior knowledge. In the present study, this was accomplished by selecting a contemporary political issue currently in the news yet a topic about which participants were unlikely to have extensive prior knowledge—a hypothesis upheld by the prior knowledge data. While students may have had some prior exposure to the topic, it was expected that they would have more limited engagement with it as a *controversial event*, in the way set up by the assigned task.

Second, researchers selected naturally occurring source material, typical of what learners would encounter through the course of conducting research on a particular topic (Strømsø et al., 2011). The importance of selecting such texts has been particularly emphasized in history studies (Britt & Aglinskias, 2002) where these documents serve as historical evidence. In this study, texts selected were similarly naturally occurring. Selecting a contemporary political issue as a topic resulted in the availability of a great deal of naturally occurring source material, which stands in contrast to topics in the more distant past about which primary source material may be limited and culled over time by historians. Third, Strømsø, Bråten, and Britt (2011) selected texts to represent a variety of perspectives on a topic, and fourth, texts were selected to have strong social implications and garner interest and engagement from learners. In the present study, both the selected topic and the task formulation (i.e., explicitly asking participants to choose between either Morsi or el-Sisi) ensured that a variety of conflicting viewpoints could be introduced to students. Selecting a topic addressing a contemporary political issue, one with real-world consequences, was intended to garner learner interest and engagement.

In addition to the above four guidelines a number of additional principles were observed in selecting sources for the present study. First, I ensured that all texts included in the source library were relevant to the target prompt. Source relevance was held constant in this study as it has been found to be a driving consideration in students' source use and evaluation (e.g., Anmarkrud et al., 2013; Rouet & Britt, 2011) and the purpose of this investigation was to examine students' source selections and judgments along other, non-relevance based dimensions (e.g., trustworthiness, interest). Source relevance was confirmed by making sure that all texts included in the document set had keywords and information pertinent to the prompt. Information from any of the included texts could have been used in forming a view on the target issue and composing a justified response.

Texts were also chosen to have a readability level appropriate for undergraduate students. Additionally, sources were selected to represent a variety of source types that students might typically use to research the given prompt. Similar to typical studies in the domain of history, texts represented varying degrees of authors' expertise and differed in intentionality, with some sources intending to present unbiased information and other sources arguing a strong point-of-view. Texts also differed in their information source, including participant accounts (i.e., primary sources) and secondary summaries and analyses (i.e., secondary sources). While all selected texts included information students could draw on in responding to the prompt, this information differed in its reliability and reputability. Across texts, information was complementary, allowing for corroboration, *and* discrepant, requiring reconciliation. It was expected that students would have to access multiple sources in order to develop a comprehensive, elaborative,

and well-reasoned response to the prompt. Table 3 presents document information about each of the texts included as a part of the task. Appendix I presents each of the six texts.

Text number. The number of sources presented to students in studies of multiple source use has varied, ranging from two (Wolfe & Goldman, 2005) to 15 texts (Stadtler & Bromme, 2007). The majority of studies have provided students with six to eight sources (e.g., Bråten et al., 2011; Rouet et al., 1996; Wiley et al., 2009). This number of texts is considered to offer a balance between providing students with a sufficient variety of sources to draw on and including a manageable amount of variation between sources for researchers to analyze. In this study, six sources were provided to students.

The lower bound of the six-to-eight range was selected because, while preserving sufficient source variation, I was interested in providing students with as few texts as possible. This is because students were able to select which sources to access and which to skip, and they were not required to visit all of the sites. Providing students with fewer texts potentially encouraged accessing of more sources, as accessing each of the texts during the research phase of the study became more viable as the number of potential sources decreased. Although this methodology was a deviation from studies that require students to access each source in a library (e.g., Bråten et al., 2011; Rouet et al., 1996), allowing for source selection and therefore students' own determinations of information need was considered to be theoretically important in creating a multiple source task.

Indeed, source selection is considered to be the initiating step in the multiple source use process and is tied to students' determinations of information need or their reasons for accessing sources (Rouet, 2006; Rouet & Britt, 2011). Likewise, an important step in the multiple source use process is students' recognition that their

information needs have been met, resulting in the cessation of source use; such determinations may have been compromised if learners were required to access all sources in the library. Allowing students to self-determine when accessing additional sources was no longer required was a necessary procedure to adopt as in this study participants were asked to make a deliberate choice about when to move from the research phase to the response phase.

Text length. The texts provided to students have been brief: in the source library used by Bråten, Strømsø, and Britt (2009) sources averaged 286 words; likewise, the historical documents introduced as part of Wineburg's study (1991) were on average 251 words in length; in a study by Wiley et al. (2009), texts explaining a complex scientific concept ranged in length from 300-500 words. The sources included in the present study were shortened to fall within this range; all texts included were under 500 words.

Text presentation. There has been a great deal of variation in the ways that texts have been presented to students and in the construction of source libraries. Modes of presentation have ranged in complexity from folders including print versions of all texts (e.g., Bråten et al., 2011) to website interfaces presenting sources to students via an electronic library (Stadtler & Bromme, 2007). For instance, Britt and Aglinskas (2002) in constructing the Sourcer's Apprentice, a digital learning environment, presented each text to students as if it were a book on a bookshelf in a library. Sources were ordered from most general to most specific, starting with a textbook source on the left and specific historical documents positioned on the right.

The first two pages of each "book" contained document information, such as what may be found on the inside of a book-jacket (e.g., author credentials, publisher). Other studies

presenting texts digitally have created mock search engine results pages as libraries from which students can access sources (e.g., Mason et al., 2010b).

Neither of these models of digital source presentation was adopted in the present study. A search engine-type interface was not used because when presenting sources as search results students have been found to strongly favor those presented at the top of the hits list (i.e., appearing first, second, or third) and consider them to be most relevant (e.g., Kammerer & Gerjets, 2012b; Mason et al., 2010b). I wanted students to consider all sources in the library to an equal extent. Further, the library-type interface, as created by Britt and Aglinskas (2002), was not favored; as the texts I was presenting to students were digital and so a more Web-based appearance was desired.

Sources were presented to students as an array. Students were able to access any source by clicking on a hyperlinked button. This format of source presentation offered a number of advantages. In particular, presenting texts as an array removed any effects arising from a list-based ordering. Kammerer and Gerjets (2010) have found that presenting sources via a grid, rather than a list, encourages more deliberative source selection. The buttons were arrayed in two rows and button order was randomly arranged for each participant. A hyperlinked array served as a middle-ground between the source presentation approaches introduced by Mason et al. (2010b) and Britt and Aglinskas (2002). Sources were presented to students via a digital library, interrupting students' conceptualization of the sources as search engine results, but were not formatted to resemble a library per-se, preserving the digital feel of the task.

Table 3

Document Information for Texts in the Multiple Source Use Task

Title	Source Type	Author/ Editor	Credentials	Publisher	Date	Location of Publication	Parent URL	Reading Level	Word Count
El Sisi's Islamist Agenda For Egypt: The General's Radical Political Vision	Analysis Essay	Robert Springborg	Professor of national security affairs at the Naval Postgraduate School	Council on Foreign Relations	July 28, 2013	Washington, DC	www.foreignaffairs.com/articles/	36.7/13.4	497
A Message to the Muslim Brotherhood: Enough Blood, Enough Lies	Blog Post	The Idealist	I am just Egyptian girl who lives in the present with the glories of the past and hopes in a better future for herself and for her country	Wordpress	Posted July 27, 2013	Alexandria, Egypt	https://theidealists11.wordpress.com/	53.1/9.6	399
Egypt Arrests 11 Islamists for Facebook Activity	Newspaper Article	Maggie Michael	Associated Press Senior Reporter, Cairo Bureau	Associated Press	January 30, 2014	Cairo, Egypt	http://bigstory.ap.org/	36/13.1	489
Egyptian Attitudes	Public Opinion Survey	Dr. James Zogby and associates	Dr. James Zogby is founder and president of the Arab American Institute, a Washington, D.C. based organization, that serves as the political and policy research arm of the Arab American community.	Zogby Research Services for Arab American Institute	September, 2013	Abu Dhabi, United Arab Emirates	www.zogbyresearchservices.com/egyptian-attitudes-2013/	44.2/13.8	446
N/A	Twitter	@Ikhwanweb	Official English language Twitter account of the Muslim Brotherhood	Twitter	February 14 – February 17, 2014	London, UK	https://twitter.com/Ikhwanweb	34.4/11.9	352
2013 Egyptian coup d'état	Wikipedia Entry	Omar Othman 95, Greyshark09, and Alhanuty	Omar Othman 95: I am Omar Othman, an Egyptian high school student from Cairo. GreyShark09: This user has created 182 articles. This user is Mediterranean. Alhanuty: This user was born and lives in the United States. This user is of Egyptian ancestry. This user is a member of WikiProject Egypt.	Wikipedia	Current version modified: February 19, 2014	N/A	http://en.wikipedia.org/	33.8/13.5	491

Note: Reading level refers to Flesh-Kincaid Readability/Flesch Grade Level

Studies have also ranged in the type and amount of information made available about sources, prior to selection. In some studies all of the sources have been directly presented with no interceding digital library. As a result students have not been provided with the opportunity to “select” sources (Bråten et al., 2011; Wineburg, 1991). Other studies have allowed students to choose sources from a library based on content (i.e., title, key-words, and brief description of site content, Mason et al., 2010b) or basic document information (i.e., author name and credentials, document type, and date, Rouet et al., 1997). In this study, I was interested in collecting behavioral data about whether or not students elected to access document information (i.e., author, publisher, title, date and location of publication, and URL) about each source they used. Consequently, only limited information about each source was provided in the digital library, prior to students selecting a text. Sources were listed in the library only by source type (e.g., newspaper article, blog), with no additional document information offered. It was hoped that requiring texts to be selected based on source type would encourage students to access additional document information about each source. Source type was considered to be a salient document feature upon which students could base their source selections, as studies have shown students to evaluate texts based on source-type based schema (Bråten et al., 2011). Further, source type was a particularly apt feature along which to organize texts in the digital library as the texts in this study were, in part, selected to present an incongruous mapping between source and information reliability. In other words, given the study topic, source type was a particularly interesting feature to use in presenting texts to students, as source types considered to be low in reliability could

nonetheless provide accurate, first-person information. Appendix J includes a screenshot of the source library.

As in other studies, texts were presented to students absent any hyperlinks to sources external to the study interface. Sources were made uniform in formatting without any graphics or advertisements (e.g., Bråten et al., 2011; Mason et al., 2010b; Wineburg, 1991).

Source use data. As students engaged in multiple source use—accessing texts in the library—log data were collected about which sources they accessed and duration of access. When sources were re-visited (i.e., accessed multiple times), time on texts was added across source visits. Based on this information, two metrics of source use were determined: the total number of sources students accessed, and the total time learners devoted to source use.

Accessing document information. When students accessed sources from the library, they were presented with just the text of each source. At the top of each source page, there was a button, “Click here to learn more about this source.” Clicking the button led students to document information about each source (i.e., title, author, publisher, date and location of publication, URL). This document information was consistent with information presented in other studies (e.g., Bråten et al., 2009; Bråten, et al., 2011; Rouet et al., 1996; Wineburg, 1991). Attending to document or source information has been found to be associated with more expert, epistemically oriented, and metacognitively engaged source use (e.g., Britt & Aglinskias, 2002; Stadtler & Bromme, 2007, 2008; Wineburg, 1991). Whether or not students elected to access document information for each source they used was recorded. Percentage of document

information accessed, computed as the ratio of the number of sources for which students accessed document information to the total number of sources they accessed, was used as a behavior-based metric of sourcing in this study. When students re-visited sources, accessing document information on any of their source visits was counted as their having accessed document information for that particular source. Appendix K includes screenshots of accessible document information.

Source evaluation. The bottom of each source page featured a button, “Back to Library,” that students could click to indicate that they have completed using a particular source. When students clicked this button, an evaluation window appeared. Students were asked to rate each source accessed along four dimensions: *trustworthiness*, *usefulness*, *interestingness*, and *information accuracy*. The final dimension (i.e., information accuracy) is not examined in the current study as, given the sample’s limited prior knowledge, it was difficult to ascertain the extent to which students may have been able to judge information accuracy.

The trustworthiness dimension has been used widely in prior research as the general dimension along which students evaluate sources based on document information (e.g., author, publisher, Bråten et al., 2009; Bråten et al., 2011). As such, the trustworthiness dimension was considered to be an epistemic dimension of source evaluation. The usefulness dimension was included as a non-epistemic source evaluation dimension representing the instrumental value of a source for meeting task goals. Given that the task or goal motivating source use has been found to be key in the multiple source use process (Rouet, 2006), this dimension was theoretically important to include and reflected students’ assessments of different sources’ functionality in responding to

the prompt. These two evaluative dimensions have been used in prior work (Rouet et al., 1996) and Afflerbach and Cho (2009) have identified assessments of text trustworthiness and usefulness as important judgments for students to make when engaged in multiple text evaluation. These two dimensions are explored in this study, as they were associated with learner characteristics and task performance.

Given the role that interest has been found to play in students' multiple source use (e.g., Bråten & Strømsø, 2006) and in reading more generally (e.g., Alexander et al., 1994; Hidi, 2001; Ozgungor & Guthrie, 2004; Schiefele & Krapp, 1996), students were also asked to rate the interestingness of each text they accessed. While prior research has focused on assessing students' general topic interest (Boscolo & Mason, 2003; Bråten & Strømsø, 2006; Gil et al., 2010b), more recent work has examined situational interest emerging during multiple source use (e.g., Bråten et al., 2014). Interest has been described as an engagement factor particularly important to multiple source use as it provides a motivational resource for students to draw on when confronted with complex tasks (e.g., multiple source use). More simply, interest may afford students the *cognitive energy* and support the attention and expending of effort necessary to engage in the complexity presented by multiple texts (e.g., Bråten & Strømsø, 2006; van Oostendorp & de Mul, 1996). In the present study, students were asked to rate each text's interestingness following source use, accessing situational aspects of interest.

Students' ratings of each source used were recorded. Learners were required to evaluate each source accessed before they were able to return to the library and proceed with the study. Appendix L includes a screenshot of the source rating window students saw after using each source.

In analyses, indices based on students' evaluations of each source used were computed. Source usefulness and interest ratings were averaged across all of the sources a particular student accessed. If students revisited sources, their ratings across visits were averaged. The use of an overall source usefulness and source interest score was considered to be an appropriate summative measure of students' overall perceptions of the instrumental utility of texts accessed in meeting task demands and engagement with sources. Because trustworthiness ratings are determined based on document information (Britt et al., 1999), these were considered to be too text-specific to be meaningfully averaged. Rather a *trustworthiness discrimination* score was computed as the difference between students' ratings of the source they found to be most trustworthy and the source they found to be least trustworthy. Discrimination in ratings has been used in prior research as a summative score of students' multiple source evaluations (e.g., Braasch, Bråten, Strømsø, & Anmarkrud, 2014; Braasch et al., 2009; Wiley et al., 2009) and Goldman et al. (2012) have found students who learn more from multiple source tasks to be more discriminant in their source evaluations. Given that, in this study, texts were purposefully selected to vary in source type, author expertise, and credibility, trustworthiness discrimination was considered to be a particularly apt and task-specific metric to adopt.

Summative measures of source evaluation, rather than six trustworthiness ratings, six usefulness ratings, six interestingness ratings (i.e., three corresponding to each source) were used for several reasons. First, using source-specific ratings would have added a very large number of predictors to the model, taxing sample size, presenting issues of multicollinearity, and offering challenges in interpretation. More importantly, due to

listwise deletion in multiple regression, those students not accessing all six sources, and therefore not having evaluation scores for texts not accessed, would have been deleted from analyses. Removing students not using all six sources from the data set would have presented validity issues associated with the restricted sample. Further the total number of sources used as a predictor would have been superfluous (i.e., all students would have accessed six sources).

Response composition. Once students determined that they had completed the research phase of the study, they were able to access the response page by clicking the, “Click here to compose your response” button in the source library. During the response phase students were first asked to position themselves in reference to the prompt (i.e., in support of either Mohamed Morsi or General el-Sisi) and then to provide specific arguments and evidence to support their position. Students also had the option of designating their position as “Other.”

Response coding. Students’ responses consisted of their designation of a position on the target issue (i.e., endorsement of Morsi, el-Sisi, or an Other option) as well as justifications for this position.

First, students’ responses were identified as indicating a support for Mohamed Morsi, General el-Sisi, or an Other option. Next, students’ justifications for their positions were coded using four different indicators. Students’ open-ended responses were coded based on (a) *word count*, (b) number of *arguments* generated, (c) *SOLO scores*, and (d) number of *citations* included.

Word count. Although word count was not a measure of response quality per se, response length has been used as an outcome measure in prior research (Burstein,

Chodorow, & Leacock, 2004; Wiley & Voss, 1999) and was included in this study as an indicator of the elaborateness of students' responses. Further, this measure was included, since students have been found to self-evaluate open-ended responses based on length.

Argument. The number of *arguments* included was a count of the number of distinct reasons participants provided in justifying their position. An argument consisted of a claim and a justification, or supporting evidence or reasoning for said claim (Brockriede, 1992; Rouet et al., 1996). For example, a single argument consisted of a student who made the claim, "Morsi was ousted by Sisi so that Sisi could gain power," and provided the following evidence, "Although initially he [el-Sisi] stated that he was not interested in holding political office, later he showed that he was interested in becoming the leader and did not show much enthusiasm for democracy." Number of arguments included is a common measure that has been used in the multiple source use literature to assess students' open-ended responses (e.g., Rouet et al., 1997; Rouet et al., 1996). Rouet et al. (1996) specify that arguments are the essay segmentation method to use when students' essays contain wide variation in statements including facts, evaluations, and claims. Interrater agreement was computed based on two rates scoring 20.81% of students' responses (n=41) and high correspondence in the number of arguments identified was found: $r(41)=0.89$. Disagreements were resolved through discussion and I coded the remainder of responses.

SOLO taxonomy. In addition to counting the number of distinct arguments students included in their responses, participants' responses were assigned a score on a modified Structure of the Observed Learning Outcome taxonomy (SOLO, Biggs & Collis 1982). The SOLO taxonomy considers not only the number of reasons students included

in their responses, but also the quality of these responses and the level of sophistication in reasoning they reflected. The SOLO taxonomy has four levels. A one on the SOLO taxonomy corresponds to a *unistructural* response, containing only a single piece of information. Meriting a two on the SOLO taxonomy corresponds to a *multistructural* response, wherein a number of reasons are identified and introduced; however, these elements are not integrated. Responses scoring a three on the SOLO taxonomy may be considered to be *relational*, or contain multiple elements or reasons that are also connected to one another or integrated. Finally, a four on the SOLO taxonomy represents a response marked by *extended abstraction* that not only presents multiple, integrated pieces of evidence but also generates generalized principles or extends responses beyond the immediate context of the problem.

In my adaption of the SOLO taxonomy, responses that evaluated competing arguments in reference to one another before arriving at a conclusion were also assigned a four. Such evaluative or refutational responses were considered to map or apply arguments in favor of one position onto another and visa-versa. For instance, responses such as, "...therefore, neither side is completely to blame for the violence. While the Muslim Brotherhood should not have turned to violence in order to regain power for Ex-President Morsi, General El-Sisi should not have ordered arrests and assaults for things posted on social media..." was assigned a four on the SOLO taxonomy because students were comparing and considering the positions of Morsi and el-Sisi in reference to one other and evaluating the merits of each before arriving at a conclusion. Put another way, consistent with an extended abstraction response, these justifications extended beyond arguments for a single position to consider both perspectives. Assigning evaluative

responses a four on the SOLO taxonomy was necessary to separate participants' essays that were integrated but only one sided (e.g., in support of Morsi) from those that considered both positions (i.e., both Morsi's and el-Sisi's). It should be noted that few participants constructed generalized responses, as would traditionally be assigned a 4 on the SOLO taxonomy, thus the taxonomy was modified to reflect and differentiate the range of responses participants produced. Generating a relational or evaluative response was thought to reflect students' engagement in core, high-level multiple source use processes (Britt et al., 1999; Rouet, 2006; Rouet & Britt, 2011). Specifically, while developing a connected written response (SOLO score 3) was considered indicative of the integration of multiple texts, an evaluative response (SOLO score 4) was seen as predicated on the corroboration and reconciliation of multiple texts.

The SOLO taxonomy was further modified by the inclusion of half points to allow a more graduated response scoring. Justifications were first assigned a score based on the type of response participants were striving for and reduced half a point if the response did not adequately reflect that SOLO level. For example, a response such as, "I have come to the conclusion that General el Sisi is somewhat [sic.] of a tyrant who needs to be stopped," was assigned a 0.5 on the SOLO taxonomy. While this student was trying to provide a single argument, corresponding to a unistructural response, this was not fully articulated and therefore the SOLO score was reduced by half a point. Interrater agreement was computed based on 13.20% of responses (n=26). Cronbach's alpha was 0.96, considered to indicate *very good* reliability (Cohen, 1960). Although perfect agreement between rates was only 59.46%, deviations in assigned SOLO scores were

small (i.e., discrepancies: $M=0.26$, $SD=0.48$) and were resolved through discussion. I coded the remainder of responses.

Citation. The final metric used was the number of citations included in students' responses. Citations were any in-text or parenthetical reference to sources in the library, appearing in students' written responses. Citations included instances of source identification based on any document information including source type, author, or title. For instance, "according to an analysis essay by Robert Springborg [sic.]" and "(Zogby and Associates, 2013)" were both considered to be citations. Citations have been considered a great deal in scoring open-ended responses, since they indicate attendance to document information and adherence to disciplinary conventions for writing (Britt & Aglinskas, 2002; Rouet et al., 1996; Strømsø & Bråten, 2014). Students were not explicitly instructed to include citations in their responses, so the number of citations included represents students' spontaneous sourcing (Strømsø, Bråten, Britt, & Ferguson, 2013). Strømsø et al. (2013) point to the importance of considering spontaneous sourcing, previously examined only to a limited extent, as a more ecologically valid examination of students' engagement with document information. In this study, citations are considered to be a manifestation of learners' sourcing skills not only during text processing but in writing as well. Further, students' association of content in texts to document information, as reflected in citation use, was considered to indicate document model construction (Britt et al., 1999). Sample response codings are presented in Appendix M.

Post task measures. After students completed their responses and clicked the "submit" button, they were taken to a final source evaluation page. Students were asked

to rank those sources they accessed during the research phases with regard to their *trustworthiness, usefulness, interestingness, and the accuracy of the information* they included. In ranking sources along each of these dimensions, students were further asked to provide an open-ended justification for each ranking in a textbox. A screenshot of the ranking and justification page is included as Appendix N. However, students' source rankings were not considered in present analyses.

Following the evaluation task, learners were asked to report their *task engagement*. A five item engagement scale was created, asking students to report how interested they were in completing the task as well as how difficult they considered the task to be and how much effort they devoted to task completion. The task engagement scale and item-specific reliability statistics are included in Appendix O. Scale if item deleted reliability analysis determined that the item asking participants to report task difficulty should be excluded. The four items scale had a Cronbach's alpha of 0.81. Students responded to each item on a seven point Likert scale.

Data Analysis Plan

Table 4 includes a listing of all variables examined as well as scale information. In Session 1 of the study, three learner related variables were assessed. Specifically, learners reported their (a) prior topic knowledge, (b) attitudes, and (c) domain general source evaluation behaviors. In Session 2 of the study, a variety of measures associated with students' multiple

Table 4

Variables Examined

	Number of Items	Reliability	Scoring	Scale
Independent Variables				
<i>Learner Characteristics</i>				
Prior Knowledge	7	$\alpha=0.91$	Binary: Correct/Incorrect	0-7
Stance	1		Rating Scale	1-7
Source Eval. Behaviors	9	$\alpha=0.86$	Rating Scale	1-7
<i>Source Use Behaviors</i>				
Total Sources			Summary of all sources accessed	0-6
Total Time			Sum of time on texts	
Percent Doc Info Accessed			Ratio of number of sources for which doc. info was accessed to all sources accessed	0-1.00
Trust Discrimination			Difference between most trustworthy and least trustworthy source ratings	0-100
Avg. Usefulness			Avg. usefulness ratings of all sources accessed	0-100
Avg. Interestingness			Avg. interestingness ratings of all sources accessed	0-100
Dependent Variables				
Word Count			Totaling of words in response	
Argument		$r=0.89$	Count of arguments	
SOLO Score		$\alpha =0.96$	Assigned a value of 0-4, with half points	0-4
Citation			Count of direct references	
Post-Hoc Measures				
Task Engagement	4	$\alpha=0.81$	Rating Scale	1-7

Note: Evaluation metrics were computed based only on those sources students accessed.

source use behaviors were collected. Specifically, log data were used to determine (a) the total number of sources learners access, (b) duration of source access, and (c) the percentage of source for which document information was accessed. During source use, learners were also asked to evaluate sources. Evaluation metrics included students' (a) discrimination in trustworthiness ratings as well as ratings of (b) average source usefulness and (c) average source interestingness, computed across all texts a particular student accessed. These learner characteristics and multiple source use behavioral variables were considered descriptively, in associating with one another, and as predictors of open-ended response quality. Four key outcome variables were examined. Specifically, students' open-ended responses were coded according to: (a) word count, (b) number of arguments included, (c) SOLO scores, and (d) number of citations provided.

To answer the first research question (i.e., What is the nature of students' multiple source use when responding to a multiple source use task?), I examined descriptive information of source use behaviors and response quality. Further, one-way analysis of variance was used to determine whether source use behaviors (e.g., time on text, source evaluations) and task performance (e.g., word count) differed by source type. Chi-squared tests were used to determine whether there was an association between source type and whether or not students elected to access document information about each text. Finally, the post-hoc task engagement measure was analyzed as associated with learner characteristics, multiple source use behaviors, and task performance.

In responding to the second research question (i.e., To what extent do source use behaviors predict response quality when students complete a multiple source task?), I

conducted four multiple regressions. Source use behaviors (i.e., number of sources, time on texts, percentage of document information accessed, discrimination in trustworthiness ratings, average usefulness rating, average interestingness rating) were used to predict each of the four target outcome variables associated with open-ended response quality (i.e., word count, number of arguments, SOLO score, number of citations).

Question three (i.e., What is the nature of the association between individual difference factors and students' manifest source use behaviors?) was answered with a series of correlations examining the bivariate relations between individual difference factors (i.e., prior knowledge, stance, and domain general source evaluation behaviors) and multiple source use behaviors (i.e., total number of sources access, total time on texts, percentage of document information accessed, discrimination in trustworthiness ratings, as well as average usefulness and interestingness ratings).

To answer the fourth research question, (i.e., To what extent do individual difference factors predict response quality when students complete a multiple source task?), I ran four multiple regressions. In each regression, learner characteristics (i.e., prior knowledge, topic stance, domain general source evaluation behaviors) were used to predict each of the four response outcomes (i.e., word count, number of arguments, SOLO scores, and citations).

I addressed the fifth and final research question (i.e., To what extent do individual difference factors) and multiple source use behaviors predict response quality when students complete a multiple source task?) by running four multiple regression models with both learner characteristics (i.e., prior knowledge, topic stance, source evaluation behaviors) and multiple source use behaviors (i.e., total number of sources access, total

time on texts, percentage of document information accessed, discrimination in trustworthiness ratings, as well as average usefulness and interestingness ratings) predicting each of the four target outcome variables (i.e., word count, number of arguments, SOLO scores, citations). These final, full models including learner characteristics and multiple source use behaviors as predictive of open-ended response quality were the central focus of this study.

Multiple Regression

The predictive models of open-ended response quality were tested using multiple regression analysis. Multiple regression was the method selected as it is a ubiquitous analytic approach in the social sciences, widely used to predict a target outcome variable based on a collection of independent predictors (Kelley & Maxwell, 2010). In particular, multiple regression has been praised for its flexibility and allowance for model complexity (Cohen, Cohen, West, & Aiken, 2002; Tabachnick & Fidell, 2001). A particular benefit of multiple regression is that it allows for the evaluation of the effects of both a collection of predictors and of particular independent variables on the dependent variable, holding constant other predictors in the model.

Multiple regression is considered to be a procedure appropriate for two types of functions: it may be used either to predict a pragmatic outcome or to explain a phenomenon (Cohen et al., 2002; Pedhazur, 1997). In the present study explanation was the goal. In particular, I was interested in the extent to which students' response quality may be explained by learner characteristics and by multiple source use behaviors, rather than in predicting students' specific levels of performance. In addition to the distinction drawn between predictive and explanatory multiple regression, multiple regression may

also be *confirmatory* or *exploratory* in nature (Kelley & Maxwell, 2010). A confirmatory approach, as is adopted in this study, is marked by the presence of well-defined research questions and theoretically justifiable variable inclusion. An exploratory approach is more diffuse, with less concrete research questions and variable definitions (Kelley & Maxwell, 2010).

The analysis used in the present study was *hierarchical* in nature. In hierarchical linear regression, in addition to having theory-driven, researcher specification of variables, the researcher also selects the order of variable entry into the model based on theoretical considerations (Kelley & Maxwell, 2010). I selected the order of variable entry to correspond to the parameters of the MD-TRACE model. The MD-TRACE model suggests that learners bring internal cognitive resources to the multiple source task (i.e., learner factors entered at Step 1), before engaging in MSU behaviors (i.e., multiple source use behaviors entered at Step 2). In the case of hierarchical regression, it is recommended that the change in R^2 be examined between nested models to ascertain whether blocks of variables entered in later stages provide significant improvements in terms of model fit relative to the more simple models. In analyses, tests of significance in R^2 change were performed based on blocks of variables entered, rather than based on specific predictors; however, the effects of specific predictors in the model (e.g., prior knowledge, time on texts) were also be considered.

In the present study, multiple regression offered particular benefits. Generally, multiple regression is robust to violations of assumptions (Tabachnick & Fidell, 2001). When assumptions are violated, estimates may still be unbiased but relations attenuated. Further, multiple regression makes no distributional assumptions with regards to the

normality of either the predictors or the outcome variable (Kelley & Maxwell, 2010; Tabachnick & Fidell, 2001). This is advantageous as many of our predictors (e.g., number of sources) and outcomes variables were skewed. Rather, normality is only considered with regard to the distribution of residuals or standardized errors terms. Further, violations of normality in the residuals may be mitigated with increased sample size (Tabachnick & Fidell, 2001).

Regression vis-a-vis Path Analysis

A form of multiple regression, measured variable path analysis, was considered as an additional analysis strategy. An advantage of path analysis is that it allows for both direct and indirect effects of predictor variables on the dependent variable to be modeled. Findings from multiple regression were scrutinized to determine whether further exploration, potentially considering mediation via path analysis, was needed. As per the theoretical framework, I was primarily interested in examining the direct effects of learner characteristics and multiple source use behaviors on target outcome variables (i.e., multiple source use task performance), rather than on examining potential relations between independent variables and mediating pathways. Further, correlation analyses determined limited relations between individual difference factors and multiple source use behaviors suggesting a lack of mediation (Baron & Kenny, 1986). More generally, I thought that my model could be most effectively and efficiently modeled with hierarchical multiple regression.

CHAPTER 4: RESULTS

In answering the five research questions encompassing the extent to which learner characteristics (i.e., prior knowledge, stance, and domain general source evaluation behaviors) and multiple source use behaviors (i.e., total number of sources accessed, time on texts, percent of document information accessed, discrimination in trustworthiness ratings, average source usefulness, and average interestingness) are associated with one another and predictive of response quality a variety of methods were used. These included descriptive statistics, ANOVA, chi-squared tests, correlations, and multiple regressions. Descriptive statistics, ANOVA, and chi-squared were used to determine the extent to which students' source use behaviors differed across source types accessed (Research Question 1). Correlation analysis was used to examine the relation between individual difference factors and multiple source use behaviors (Research Question 3).

Multiple regression analyses were used to examine three sets of models. The first set of regressions used multiple source use behavioral variables to predict each of the outcome variables (i.e., word count, number of arguments, SOLO scores, and number of citations; Research Question 2). The second set of regression models used learner characteristics to predict each of the open-ended response metrics (Research Question 4). Finally, a set of joint models, using both learner characteristics and multiple source use behaviors to predict response outcomes, was run (Research Question 5). Model comparison was used to determine whether multiple source use behaviors explained variability in performance over and above that accounted for by learner characteristics alone.

This chapter presents results specific to each research question, in order. Data examination and assumption checking to determine the appropriateness of using each procedure are explained, as they pertain to each research question. Table 5 presents descriptive information for each variable examined in subsequent analyses.

Table 5

Descriptives of Variables Examined

	Mean	Standard Deviation	Reported Range
Independent Variables			
<i>Learner Characteristics</i>			
Prior Knowledge	2.32	2.61	0-7
Stance	2.05	1.54	1-7
Source Evaluation Behaviors	4.45	1.22	1-7
<i>Source Use Behaviors</i>			
Total Sources	4.68	1.54	0-6
Total Time	15.46min	9.22 min	0-78.46 min
Percent Doc Info Accessed	0.44	0.42	0-1.00
Trust Discrimination	54.16	26.77	0-100
Avg. Usefulness	60.64	15.50	4-100
Avg. Interestingness	58.67	17.55	0-93.33
Dependent Variables			
Word Count	188.22	133.28	0-745
Argument	4.75	3.00	0-16
SOLO Score	2.75	1.12	0-4
Citation	1.46	1.74	0-7
Post-Hoc Measure			
Task Engagement	4.25	1.63	1-7

Descriptives of Multiple Source Use

Research Question 1: What is the nature of students' multiple source use when responding to a multiple source use task?

In responding to the first question, descriptive statistics are presented in five sections. First, descriptives of students' source access are introduced. Then, one-way analysis of variance and chi-squared tests are used to examine whether source use behavioral indicators (e.g., time on text) differ by source type. After that, descriptives of students' open-ended responses are presented and examined as associated with type of sources accessed. Finally, the associations between individual difference factors, multiple source use behaviors, and performance and students' reported task engagement is presented.

Multiple Source Use Behaviors

Students' used an average of 4.67 sources ($SD=1.54$). Figure 4 displays the percentage of students accessing each number of sources.

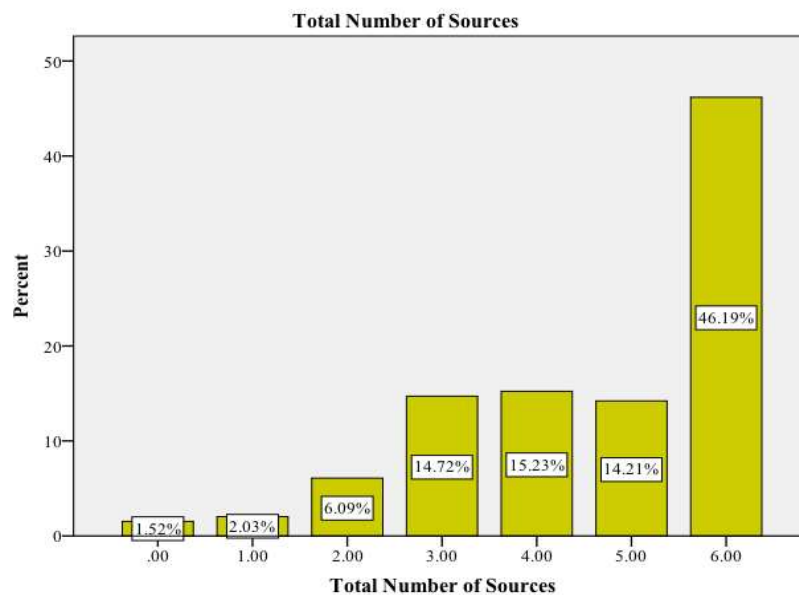


Figure 4. Frequency of each number of sources accessed.

Table 6 presents descriptives regarding the frequency with which each source was used.

Table 6

Descriptives of Students' Source Access

Source	Percentage Use (N)
Blog	61.42% (n=121)
Essay	79.19% (n=156)
Newspaper	94.42% (n=186)
Public Opinion Survey	79.19% (n=156)
Twitter	76.65% (n=151)
Wikipedia	76.65% (n=151)

Students devoted an average of 15.46 minutes to source use overall ($SD = 9.22$) and spent an average of 3.31 minutes ($SD= 2.63$) per source. In terms of source evaluations, students accessed document information for 44.28% of sources accessed ($SD=0.42$).

Multiple Source Use Behaviors By Source Type

Table 7 includes descriptive information of students' multiple source use, by source accessed. One-way analysis of variance was used to examine whether duration of access and source evaluations (i.e., trustworthiness, usefulness, and interestingness) differed by source type. Chi-squared goodness of fit was used to determine whether there was an association between source type and students accessing document information for each source.

Chi-squared determined that there was, indeed, a significant association between source type and accessing document information, $X^2(5) = 11.54, p < .05$. Examining standardized residuals (above 1.0) determined that when accessing the analysis essay students were significantly more likely to consult document information (53.85%).

Conversely, when accessing twitter (40.40%) and the Wikipedia entry (39.74%) students were less likely to consider document information.

Table 7

Descriptives of MSU by Source

Source	Percent Students Accessing	Avg. Time Spent in minutes	Percent Access Doc Info	Avg. Trustworthiness	Avg. Utility	Avg. Interest
Blog	61.42% (n=121)	<i>M</i> =2.17 <i>SD</i> =1.90	51.24% (n=62)	<i>M</i> =30.17 <i>SD</i> =22.63	<i>M</i> =48.14 <i>SD</i> =28.62	<i>M</i> =60.28 <i>SD</i> =27.23
Essay	79.19% (n=156)	<i>M</i> =4.43 <i>SD</i> =2.67	53.85% (n=84)	<i>M</i> =64.03 <i>SD</i> =24.24	<i>M</i> =67.16 <i>SD</i> =22.84	<i>M</i> =60.32 <i>SD</i> =26.41
Newspaper	94.42% (n=186)	<i>M</i> =4.65 <i>SD</i> =2.94	46.77% (n=87)	<i>M</i> =71.20 <i>SD</i> =20.37	<i>M</i> =72.61 <i>SD</i> =19.48	<i>M</i> =62.63 <i>SD</i> =24.56
Public Opinion Survey	79.19% (n=156)	<i>M</i> =2.74 <i>SD</i> =2.01	52.56% (n=82)	<i>M</i> =64.53 <i>SD</i> =22.29	<i>M</i> =60.62 <i>SD</i> =25.97	<i>M</i> =55.33 <i>SD</i> =28.58
Twitter	76.65% (n=151)	<i>M</i> =1.60 <i>SD</i> =1.37	40.40% (n=61)	<i>M</i> =31.40 <i>SD</i> =23.42	<i>M</i> =43.29 <i>SD</i> =29.05	<i>M</i> =57.86 <i>SD</i> =28.44
Wikipedia	76.65% (n=151)	<i>M</i> =3.70 <i>SD</i> =2.73	39.74% (n=60)	<i>M</i> =48.73 <i>SD</i> =24.94	<i>M</i> =67.49 <i>SD</i> =23.30	<i>M</i> =59.42 <i>SD</i> =26.23
Total		<i>M</i> =3.31 <i>SD</i> =2.63	44.28% <i>SD</i> =0.42	<i>M</i> =53.26 <i>SD</i> =27.96	<i>M</i> =60.80 <i>SD</i> =26.92	<i>M</i> =59.39 <i>SD</i> =26.89

One-way analysis of variance was used to examine differences in behavioral indicators of source use (i.e., time on text, trustworthiness, usefulness, and interestingness ratings) by source type. Although these behavioral indicators deviated from normality, see Appendix P for Q-Q plots, ANOVA has been found to be quite robust to violations of univariate normality (Keselman, Algina, Lix, Wilcox, & Deering, 2008; Schmider, Ziegler, Danay, Beyer, & Bühner, 2010). Bonferonni’s method for post-hoc comparisons with unequal sample size was used. Alpha levels were set at 0.05 and not adjusted for family-wise error. Adjustments for multiple comparisons have been critiqued as limiting

power, inflating Type II error, and as presenting ambiguity with regard to what may be classified as a “family” (Feise, 2002; O’Keefe, 2003). When comparisons are planned and theoretically justified, as in this case, alpha levels may not require adjustment (Drachman, 2012; Keppel & Zedek, 1989).

Across source types, students differed in the amount of time devoted to using each source, $F(5, 915) = 42.62, p < 0.001, \eta^2 = 0.19$, corresponding to a moderate effect size. Students spent significantly less time on the blog ($M=2.17, SD=1.90$) than on the newspaper article ($M=4.65, SD=2.94$), $p < .001$, the essay ($M=4.43, SD=2.67$), $p < .001$, and Wikipedia ($M=3.70, SD=2.73$), $p < .001$. Further, students used twitter ($M=1.60, SD=2.63$) for significantly less time than they did the newspaper article, $p < .001$, the essay, $p < .001$, the public opinion survey ($M=2.74, SD=2.01$), $p < .001$, and Wikipedia $p < .001$. The public opinion poll was accessed for significantly less time than the newspaper article, $p < .001$, the essay, $p < .001$, and Wikipedia, $p < .01$. Students examined the newspaper for significantly more time than they did Wikipedia, $p < .01$.

Students differed significantly in their evaluations of source trustworthiness, $F(5, 915) = 90.19, p < .001, \eta^2 = 0.33$, corresponding to a moderate-strong effect size. Post-hoc analyses using Bonferonni’s post-hoc analyses for unequal groups determine that students considered the blog ($M=30.17, SD=22.63$) to be significantly less trustworthy than the newspaper article ($M=71.20, SD=20.37$), $p < .001$, the essay ($M=64.03, SD=24.24$), $p < .001$, the public opinion poll ($M=64.53, SD=22.29$), $p < .001$, and Wikipedia ($M=48.73, SD=29.94$), $p < .001$. Likewise, twitter ($M=31.40, SD=23.42$) was rated as significantly less trustworthy than the newspaper article, $p < .001$, the essay, $p < .001$, the public opinion poll, $p < 0.001$, and Wikipedia, $p < .001$. Wikipedia was considered to be less trustworthy

than the newspaper, $p<.001$, the essay, $p<.001$, and the public opinion poll, $p<.001$. In summary, participants considered the blog and twitter to be significantly less trustworthy than the other documents in the source set. Wikipedia was considered less trustworthy than the analysis essay, public opinion survey, and the newspaper.

One-way analysis of variance also determined that differences in usefulness ratings differed by source type, $F(5, 915) = 34.10, p<.001, \eta^2=0.16$, corresponding to a moderate effect size. The blog ($M=48.14, SD=28.62$) was considered to be significantly less useful than the newspaper article ($M=72.61, SD=19.48$), $p<.001$, the essay ($M=67.16, SD=22.84$), $p<.001$, the public opinion poll ($M=60.62, SD=25.97$), $p<.01$, and Wikipedia ($M=60.80, SD=26.92$), $p<.001$. Likewise, twitter ($M=43.29, SD=29.05$) was considered to be significantly less useful than the newspaper, $p<.001$, the essay, $p<.001$, the public opinion survey, $p<.001$, and Wikipedia, $p<.001$. Finally, the newspaper article was considered to be significantly more useful than the public opinion poll, $p<.001$. No significant differences across sources were found in students' interestingness ratings, $F(5, 915) = 1.42, p=0.22$.

Figure 5 presents the standardized values of time spent and evaluative ratings by source type.

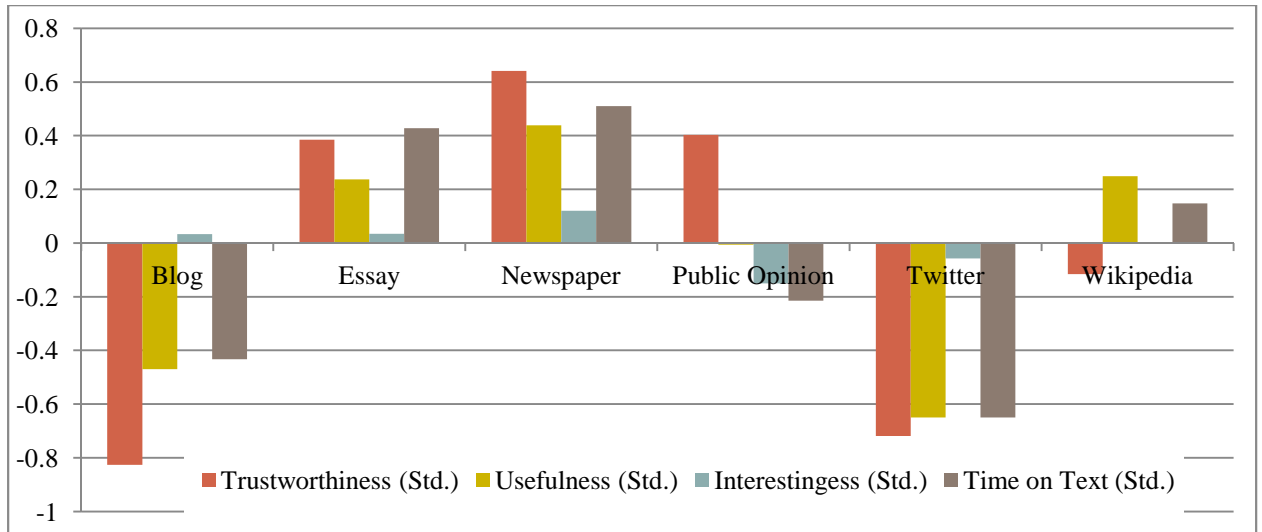


Figure 5. Standardized time spent and source evaluation ratings by source type.

Task Performance

In composing their responses, students were first asked to report whether they supported Mohamed Morsi, General el-Sisi, or an “Other” option. Mohamed Morsi was supported by 28.43% of respondents (n=56), General el-Sisi was supported by 37.06% of respondents (n=73), and the Other option was supported by 32.49% of students (n=64). Students endorsing the other option most often described the limitations associated with both leaders and advocated that neither of them should be in power. As one student, selecting the other option, explained:

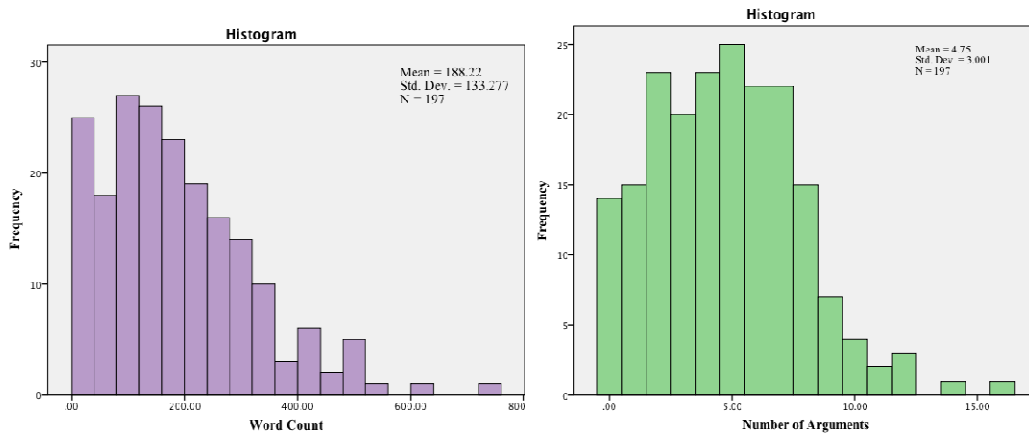
“The United States should not support General el Sisi nor Morsi and the Muslim Brotherhood. This is because these two groups impose extreme and opposite views on the public. General el Sisi focuses on a military-backed regime that strongly centers on Islamism as the end-all-be-all...The Muslim Brotherhood resort to violence and purposefully cause riots that result in the deaths or injuries of innocent by-standers.”

A chi-squared test determined that there were no significant differences in the proportion

of students selecting each option, $X^2(2) = 2.25, p=0.33$.

Prior to task engagement, participants were similarly asked to identify whether they supported Morsi, el-Sisi, and Other option, or whether they did not know. The vast majority of participants (91.67%, $n=132$) said they did not know who should be in power, with only 2.08% of students supporting Morsi ($n=3$), 3.47% supporting el-Sisi ($n=5$), and 2.78% endorsing an Other option ($n=4$). Remaining participants did not designate who should hold power in Egypt, prior to task completion. This suggests that multiple source use, at the least, was able to engender commitment to a position from students.

Students spent an average of 7 minutes and 42 seconds composing responses ($SD=6.15$ minutes) The distribution of each of the outcome variables (i.e., word count, number of arguments, SOLO scores, citations) is presented in Figure 6.



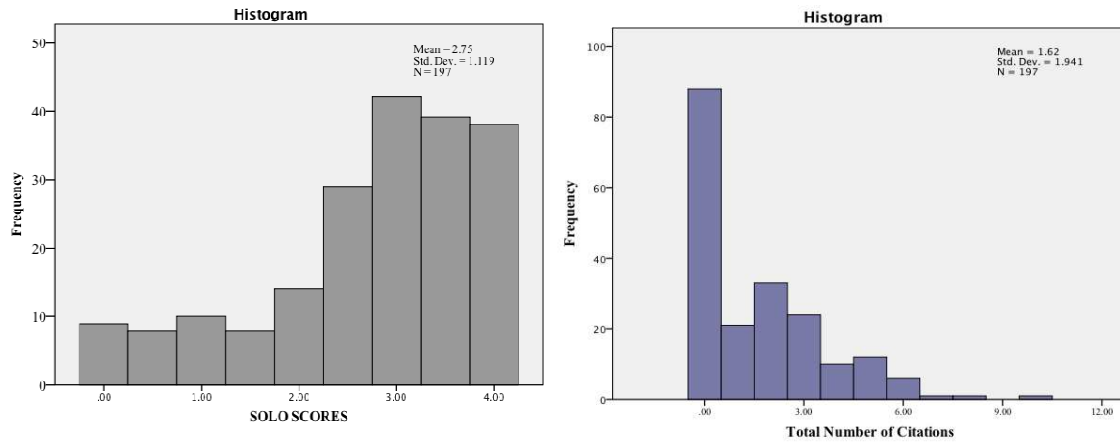


Figure 6. Histograms of each open-ended outcome variable, word count, number of arguments, SOLO scores, and number of citations.

Task Performance by Source Access

Independent sample t-tests were used to examine the relation between source types accessed (or not) and performance on the open-ended multiple source use task. In instances when the assumption of homogeneity of variance across groups was violated, as indicated by Levene’s test, adjusted degrees of freedom were used. T-test is robust to violations of normality with large samples (i.e., above 80 according to Sawilowsky & Blair, 1992; above 50 according to Ratcliffe, 1968). Table 8 presents descriptives of source access and performance.

Students accessing the blog had significantly higher SOLO scores ($M=2.91$, $SD=1.07$) than students who did not, ($M=2.48$, $SD=1.44$), $t(195) = 2.69$, $p < .01$, Cohen’s $d=0.39$. No significant differences in performance outcomes were found for students accessing the analysis essay or not, $ts(195) = 0.07-1.43$, $ps > 0.15$. Those students accessing the newspaper had a significantly higher word count ($M=192.76$, $SD=129.53$), $t(195) = 1.98$, $p < .05$, Cohen’s $d=0.28$, and significantly higher SOLO scores ($M=2.82$, $SD=1.07$) than students who did not (Wordcount: $M=111.45$, $SD=175.72$; SOLO:

$M=1.50$, $SD=1.30$), $t(195) = 3.94$, $p<.001$, Cohen's $d=0.56$. Using the public opinion survey, as compared to not, was associated with open-ended responses with significantly higher word counts, $t(195) = 2.72$, $p<.01$, Cohen's $d=0.39$, number of arguments, $t(195) = 2.22$, $p<.05$, Cohen's $d=0.32$, SOLO scores, $t(\text{adj. } 53.70) = 3.56$, $p=0.001$, Cohen's $d=0.97$, and citations included, $t(195) = 2.53$, $p=0.01$, Cohen's $d=0.36$. Using twitter corresponded to significantly more arguments ($M=4.99$, $SD=2.99$) included in responses, $t(195) = 2.06$, $p<.05$, Cohen's $d=0.30$, and significantly higher SOLO scores, ($M=2.85$, $SD=1.05$), $t(\text{adj. } 65.20) = 2.19$, $p<.05$, Cohen's $d=0.31$, as compared to not accessing twitter (Argument: $M=3.96$, $SD=2.94$; SOLO: $M=2.40$, $SD=1.26$).

Table 8

Performance Descriptives by Source Type

	Word Count		Arguments		SOLO Scores		Citations	
	Not Used	Accessed	Not Used	Accessed	Not Used	Accessed	Not Used	Accessed
Blog (n=121)	<i>M</i> =169.22 <i>SD</i> =134.53	<i>M</i> =200.15 <i>SD</i> =131.64	<i>M</i> =4.39 <i>SD</i> =3.39	<i>M</i> =4.97 <i>SD</i> =2.72	<i>M</i> =2.48** <i>SD</i> =1.44	<i>M</i> =2.91** <i>SD</i> =1.07	<i>M</i> =1.34 <i>SD</i> =1.86	<i>M</i> =1.53 <i>SD</i> =1.67
Analysis Essay (n=156)	<i>M</i> =173.93 <i>SD</i> =149.49	<i>M</i> =191.97 <i>SD</i> =128.97	<i>M</i> = 4.37 <i>SD</i> =3.37	<i>M</i> =4.85 <i>SD</i> =2.90	<i>M</i> =2.52 <i>SD</i> =1.16	<i>M</i> =2.80 <i>SD</i> =1.10	<i>M</i> =1.44 <i>SD</i> =2.04	<i>M</i> =1.46 <i>SD</i> =1.66
Newspaper (n=186)	<i>M</i> =111.45* <i>SD</i> =175.72	<i>M</i> =192.76* <i>SD</i> =129.53	<i>M</i> =3.36 <i>SD</i> =4.27	<i>M</i> =4.83 <i>SD</i> =2.90	<i>M</i> =1.50*** <i>SD</i> =1.30	<i>M</i> =2.82*** <i>SD</i> =1.07	<i>M</i> =0.73 <i>SD</i> =2.10	<i>M</i> =1.50 <i>SD</i> =1.72
Public Opinion (n=156)	<i>M</i> =138.66** <i>SD</i> =135.32	<i>M</i> =201.24** <i>SD</i> =130.06	<i>M</i> =3.83* <i>SD</i> =3.71	<i>M</i> =4.99* <i>SD</i> =2.75	<i>M</i> =2.13** <i>SD</i> =1.29	<i>M</i> =2.91** <i>SD</i> =1.01	<i>M</i> =0.85* <i>SD</i> =1.67	<i>M</i> =1.62* <i>SD</i> =1.73
Twitter (n=151)	<i>M</i> =165.61 <i>SD</i> =141.79	<i>M</i> =195.11 <i>SD</i> =130.28	<i>M</i> =3.96* <i>SD</i> =2.94	<i>M</i> =4.99* <i>SD</i> =2.99	<i>M</i> =2.40* <i>SD</i> =1.26	<i>M</i> =2.85* <i>SD</i> =1.05	<i>M</i> =1.37 <i>SD</i> =1.78	<i>M</i> =1.48 <i>SD</i> =1.74
Wikipedia (n=151)	<i>M</i> =192.93 <i>SD</i> =147.35	<i>M</i> =186.78 <i>SD</i> =129.18	<i>M</i> =4.00* <i>SD</i> =2.13	<i>M</i> =4.97* <i>SD</i> =2.19	<i>M</i> =2.76 <i>SD</i> =1.04	<i>M</i> =2.74 <i>SD</i> =1.14	<i>M</i> =1.57 <i>SD</i> =1.64	<i>M</i> =1.42 <i>SD</i> =1.78

Those selecting Wikipedia produced significantly more arguments ($M=4.97$, $SD=3.19$) than students who did not ($M=4.00$, $SD=2.13$), $t(\text{adj. } 111.95) = 2.39$, $p < .05$, Cohen's $d=0.45$. Figure 7 displays performance on the open-ended task by source accessed. All response metrics were standardized (i.e., word count, arguments, SOLO scores, citations).

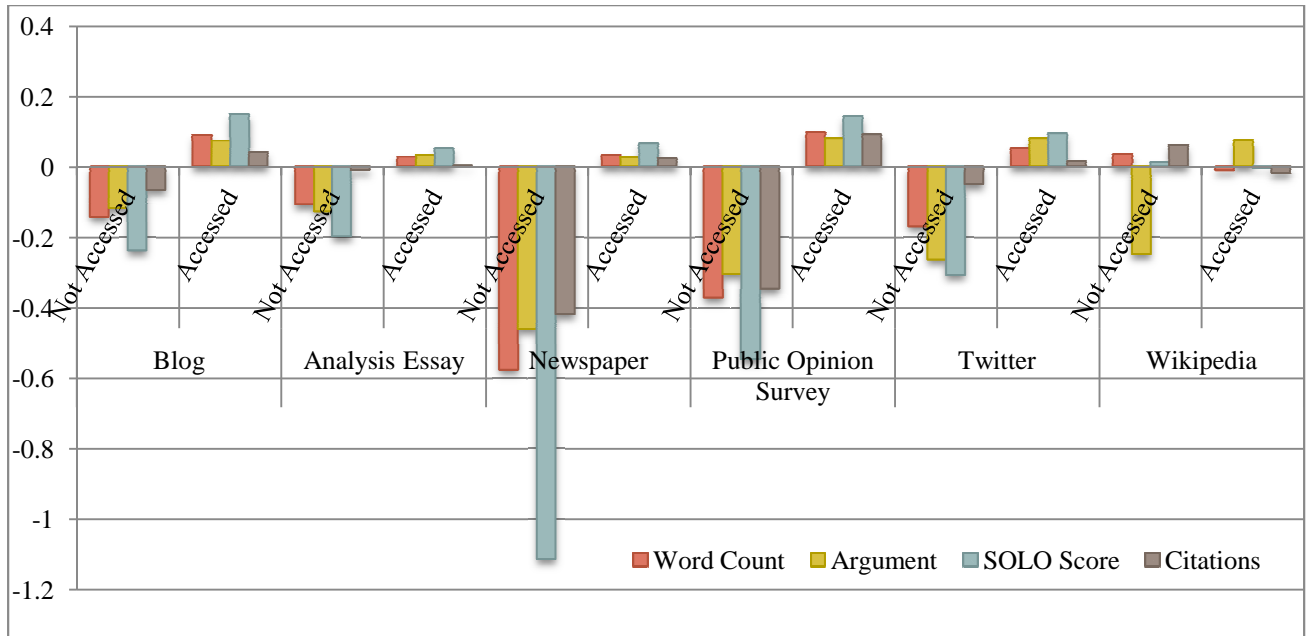


Figure 7. Performance by type of source accessed.

Figure 8 presented the percentage of students accessing each source as well as the proportion of those using each source accessing document information and citing that source in their responses.

Chi-squared was used to determine whether there was an association between source type and citing a particular source, at least once, in a generated response. Indeed, there was a significant association, $X^2(5)=71.21$, $p < 0.001$. Examining standardized residuals, determined that students were significantly less likely to cite the blog, twitter, and Wikipedia in their responses as well as more likely to cite the newspaper and the analysis essay.

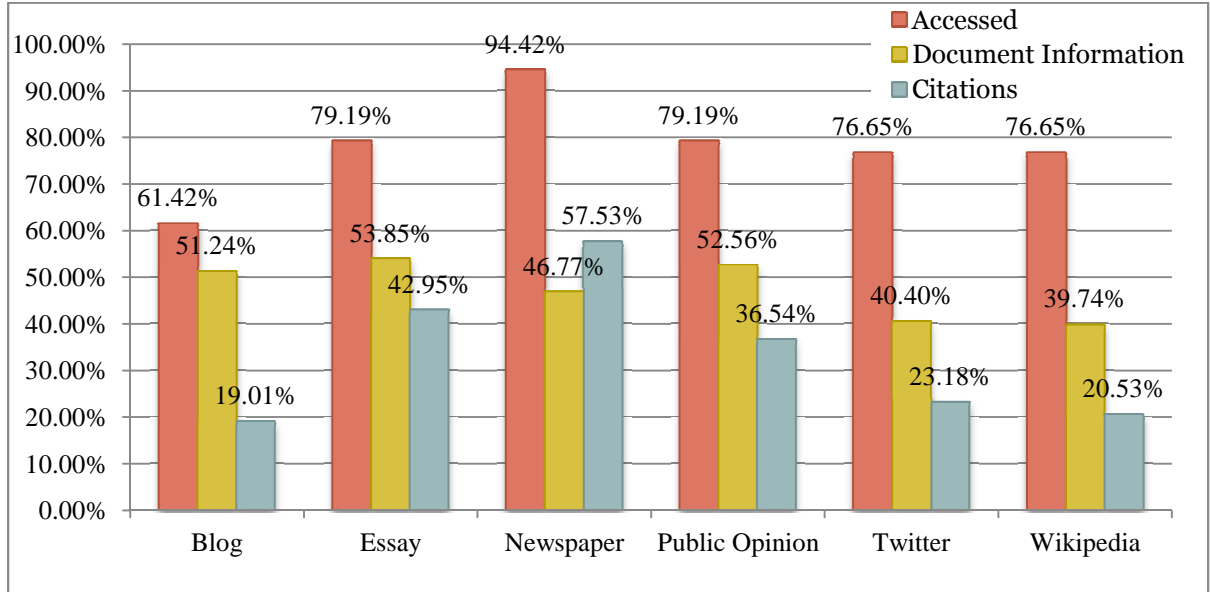


Figure 8. Access, document information use, and citation by source type.

Associations with Task Engagement

Spearman's rank-order correlation was used to examine the association between reported task engagement and learner characteristics, multiple source use behaviors, and open-ended response performance. Spearman's correlation was used as a measure of association due to violations of normality in the distributions of many of the variables, see Appendix P. Table 9 displays the relations between task engagement and individual difference factors and multiple source use behaviors. Table 10 displays the relation between task engagement and task performance.

Table 9

Correlation of Individual Difference Factors and Behavioral Variables and Task Engagement

	1	2	3	4	5	6	7	8	9	10
1. Knowledge	--									
2. Attitudes	0.63**	--								
3. Source Evaluation	0.26**	0.29**	--							
4. Total Source	-0.03	0.07	0.16*	--						
5. Total Time	-0.02	0.00	0.06	0.49**	--					
6. Accessing Doc Info	0.00	0.03	0.14*	0.25**	0.41**	--				
7. Trust Discrim.	0.07	0.12	0.08	0.46**	0.22**	0.15*	--			
8. Avg. Usefulness	0.05	0.20**	0.07	0.01	0.17*	0.16*	-0.09	--		
9. Avg. Interestiness	0.14*	0.30**	0.18*	0.07	0.21**	0.11	-0.04	0.64**	--	
10. Engagement	0.19*	0.33**	0.16*	0.08	0.12	0.14	0.18*	0.18*	0.41**	--

Task engagement was significantly associated with each of the individual difference factors: prior knowledge, $\rho(176)=0.19, p<0.05$, attitudes, $\rho(174)=0.33, p<0.001$, and domain general source evaluation behaviors, $\rho(175)=0.16, p<0.05$. Further, engagement was significantly associated with all of the source evaluation dimensions: trustworthiness, $\rho(184)=0.18, p<0.05$, usefulness, $\rho(181)=0.18, p<0.05$, and interestingness, $\rho(181)=0.41, p<0.001$.

Table 10

Correlation of Response Metrics and Task Engagement

	1	2	3	4	5
1. Word Count	--				
2. Arguments	0.74**	--			
3. SOLO Scores	0.70**	0.73**	--		
4. Citations	0.63**	0.33**	0.43*	--	
5. Engagement	0.15	0.14	0.11	0.01	--

Students' ratings of task engagement were not significantly associated with response metrics, $r_s(197) = 0.01 - 0.15, p_s \geq 0.05$.

Prior to presenting results for Research Question 2, examining the relation between multiple source use behaviors and open-ended response performance using multiple regression, as well as research questions four and five also using regression analyses, the assumptions necessary for multiple regression are examined.

Multiple Regression Assumption Checking

In multiple regression, while coefficient estimates, in and of themselves, are not based on any underlying assumptions, in order to draw inferences from coefficient estimates or to construct confidence intervals, key assumptions have to be satisfied. Despite multiple regression being generally robust to violations of assumptions,

particularly when sample size is greater than 120 (Tabachnick & Fidell, 2001), four necessary assumptions were examined and corrective measures undertaken, as needed.

There are four key assumptions associated with multiple regression. First, observations and errors should be independent and uncorrelated. Second, residuals (i.e., errors of prediction or the difference between Y expected and Y observed) should follow a normal distribution. Third, there is an assumption of homoscedasticity or that error variance is constant across all values of the dependent variable. Finally, there is an assumption of linearity or that the relation between predictors and the outcome variable is linear. Of these assumptions, the fourth, the presence of a linear relation between combined predictors and the outcome variable, is both among the most important and the most overlooked (Gelman & Hill, 2007).

While the first assumption, independence of observations, may be accomplished through design, assumptions associated with normality, homoscedasticity, and linearity may all be ascertained by examining residual plots of error terms against predicted values. Analyses of residual plots as well as additional assumption checking measures are discussed as pertaining to each assumption.

Additionally, regression has been found to be quite susceptible to outliers, requiring the examination of case-based statistics for influence and leverage. As per Tabachnick and Fidell (2001), cases with standardized residuals beyond 3 standard deviations may be considered to be outliers. Finally, in a multiple regression framework, care must be taken to avoid multicollinearity, or redundancy in predictors explaining common variance in the dependent variable. Statistics on multicollinearity for each model are considered.

Normality

The normality assumption was investigated by examining the normality probability plot (i.e., p-p plot) for each of the models. A p-p plot plots standardized residuals against their theorized normal values, with adherence to a straight line corresponding to a normal distribution of residuals. Appendix Q present residual p-p plots for models of each of the four outcome variables (i.e., word count, arguments, SOLO scores, and citations) predicted based on multiple source use behaviors (Research Question 2), learner characteristics (Research Question 4), and the full model (Research Question 5). As can be seen, residuals adequately approximated a normal distribution, with the exception of residuals for models predicting the number of citations students included in their responses. Histograms of residuals for each model are presented in Appendix R.

A recommended method for addressing violations of normality is through variable transformation (Tabachnick & Fidell, 2011). Further, logarithmic transformation, in particular, has been recommended for data that are positively skewed, as are the number of citations (Bland & Altman, 1996a; Lawner, Weinberg, Abramowitz, 2008), with the added benefit of more ready interpretability as compared to other transformation types (Bland & Altman, 1996b; Cleveland, 1984; Osborne, 2002). The number of citations students generated were transformed by a $\log(X+C)$ conversion, with $C=1$. A histogram and residual p-p plots of the transformed citation variable predicted based on multiple source use behaviors, learner characteristics, and the full model are presented in Appendix Q. The transformed citation variable is used in all subsequent analyses. However this transformation, as well as other potential transformations examined, did not

resolve all limitations in normality, thus results should be interpreted cautiously, with possible inflation of Type II error. This is particularly true for the model using learner characteristics to predict the number of citations included in students' responses (Research Question 4), whose residuals deviated from normality to the greatest extent.

Homoscedasticity and Linearity

Assumptions of homoscedasticity (i.e., constant variance of error terms) and linearity were investigated by examining residual plots for each model. Residual plots graph standardized predicted values on the dependent variable against standardized errors of prediction. Homoscedasticity assumes constant variance of error terms. In particular, residual plots can be examined to determine whether standardized residuals are uniformly distributed across value of the dependent variable. If the residuals appear to spread out along the x-axis in a fan-like pattern, this may indicate heteroscedasticity. Put another way, for the assumptions of homoscedasticity to hold, the band enclosing standardized residuals should be of uniform width across values of the dependent variable (Tabachnick & Fidell, 2001). As can be seen in Appendix Q the assumptions of homoscedasticity for each of the outcome variables appear to be preserved, especially as Tabachnick and Fidell (2001) suggest that slight heteroscedasticity has limited effects, with least squares estimation remaining unbiased (Lattin, Carroll, & Green, 2003). Heteroscedasticity is of concern when the widest spread of residuals is three times the spread of the narrowest band. In this case, I considered the assumption of homoscedasticity to be satisfied.

The linearity assumption requires a linear relation between residuals and predicted scores on the dependent variable (Tabachnick & Fidell, 2001). To demonstrate linearity, it is desirable for residuals to be randomly scattered around the 0 line, with no systematic

pattern across values. If there is a curvilinear pattern to the dispersion of the residuals, this may indicate a violation of linearity. However, regression is relatively robust to violations of linearity; estimates may simply be attenuated and relations underestimated (Tabachnick & Fidell, 2001).

In examining the residual plots, violations of the linearity assumption were of concern. Therefore partial plots, displaying bivariate relations between residuals of the outcome variable and each of the independent variables when regressed on the other predictors in the model, were examined. These are displayed in Appendix S. While slight deviations from linearity may be acceptable, the residual plots suggested a more substantial deviation from linearity. However, as can be seen by examining the partial plots, while these do not demonstrate a clear linear relation, they also fail to suggest an alternative pattern. When there is difficulty in determining the appropriate model, the linear model may function as well as any other (Prophet StatGuide, 1997). In particular, transformations introduced to compensate for violations of linearity have focused on modeling potential curvilinear relations, which also do not seem to be present (Nimon, 2012). Results from current analyses should be interpreted with caution as violations of linearity may result in an underestimation of variance explained, increasing the risk of Type II error for those independent variables having a non-linear relationship with the dependent variable and an increase in Type I error in the coefficient estimation of other independent variables that do have a linear relationship with the outcome variable (Tabachnick & Fidell, 2001).

Multicollinearity

To check for multicollinearity, two procedures were undertaken. First, a correlation matrix of all independent variables was examined to ensure that no variables were correlated above $r=0.70$, as recommended by Tabachnick & Fidell (2001). All correlations were found to be below this limit, as demonstrated in the correlation matrix presented in Table 11. Further, tolerance and the variance inflation factors for each of the models were examined. Tolerance examines the extent to which the variability in particular independent variables is not explained by other variables included in the model. A tolerance value below 0.10 is considered to indicate potential multicollinearity. The variance inflation factor (VIF) is the inverse of the tolerance and values above 10.00 are considered to indicate an issue with multicollinearity. As seen in Table 12, across the three models there did not appear to be issues stemming from multicollinearity. The final multicollinearity statistic examined was the condition index that assesses the dependency of predictors on one another. Condition indices above 30 indicate a serious problem with collinearity. Across these indices no problems with multicollinearity were identified

Outliers

Although outliers may be identified by examining a scatterplot of residuals, casewise diagnostics were used in the present study. Specifically, cases with standardized residuals greater than 3.0, as recommended by Field (2013), were considered for removal. Across models, three cases were repeatedly identified as potential outliers. Cook's distance, a score expressing the effects of one case on the regression line as a whole, was used to check if these outliers were having an undue *influence* on the regression model; a value greater than 1 would have been considered to

be concerning (Cook & Weisberg, 1982). However, across models, Cook's distance for these cases was below 1.0 (maximum: 0.09). As such, these cases were retained and believed not to have undue influence on models.

Table 11
Correlation Matrix of Independent Variables

	1	2	3	4	5	6	7	8	9
1. Knowledge	--								
2. Attitudes	0.58	--							
3. Source Evaluation Behave	0.23	0.29	--						
4. Total Source	-0.02	-0.04	0.08	--					
5. Total Time	-0.04	0.03	0.11	0.47	--				
6. Accessing Document Info	0.00	-0.01	0.14	0.41	0.25	--			
7. Trust Discrimination	0.05	0.10	0.06	0.18	0.56	0.16	--		
8. Avg. Usefulness	0.08	0.17	0.06	0.17	0.03	0.14	-0.08	--	
9. Avg. Interestiness	0.16	0.26	0.16	0.21	0.13	0.12	0.00	0.65	--

Table 12
Multicollinearity Statistics

Predictors	Model 1: Learner Characteristics			Model 2: MSU Behaviors			Model 3: Learner Characteristics and MSU Behaviors		
	Tolerance	VIF	Condition Index	Tolerance	VIF	Condition Index	Tolerance	VIF	Condition Index
Knowledge	0.66	1.52	2.81				0.65	1.54	3.49
Attitudes	0.63	1.58	4.50				0.60	1.68	4.69
Source Eval. Behaviors	0.91	1.10	10.23				0.89	1.13	6.66
Total Source				0.60	1.67	5.81	0.60	1.66	6.83
Total Time				0.69	1.45	3.93	0.69	1.45	7.48
Accessing Doc Info				0.82	1.23	6.09	0.79	1.27	11.02
Trust Discrimination				0.71	1.40	11.39	0.71	1.41	13.87
Avg. Usefulness				0.56	1.77	14.56	0.54	1.86	17.84
Avg. Interestiness				0.57	1.76	17.06	0.52	1.92	21.84

Based on validation of assumptions, multiple regression was carried out using learner characteristics, multiple source use behaviors, and their combination to predict each of the four outcome variables (i.e., word count, arguments, SOLO scores, and number of citations). Models using multiple source use behaviors to predict each of the four outcome variables are presented in response to Research Question 2. Models using learner characteristics to predict the target outcomes are presented for Research Question 4. Finally, the full model, including learner characteristic and multiple source use behaviors, to predict open-ended response performance is presented to answer Research Question 5.

Multiple Source Use Behaviors Predicting Response Quality

Research Question 2: To what extent do source use behaviors predict response quality when students complete a multiple source task?

The second research question asked: To what extent do source use behaviors (i.e., number of sources selected, time on texts, accessing of document information, and source evaluation ratings) predict response quality when students complete a multiple source task? To answer this research question, multiple regression was run predicting each of the open-ended response metrics (i.e., word count, number of arguments, SOLO score, and number of citations) based on indicators of multiple source use. Six predictors were used in the model: total number of sources accessed, time on texts, proportion of sources for which students accessed document information, discrimination in trustworthiness ratings, average usefulness ratings, and average interestingness ratings.

The full regression model predicting word count was overall significant, $F(6, 187) = 5.20, p < .001$, explaining 14.30% of the variance in word count, corresponding to a

medium effect size ($R^2 = 0.14$, Adj. $R^2=0.12$). Table 13 presents a summary of the multiple regression model predicting word count. Accessing document information and students' ratings of average source interestingness were both significant predictors in the model.

Table 13

Multiple Source Use Behaviors Predicting Word Count

Variable	B	Std. Error B	β^*	t	Sig.	Correlations		
						Zero	Partial	Part
Total Source	1.58	7.94	0.02	0.20	0.84	0.18	0.01	0.01
Total Time	0.00	0.00	0.14	1.71	0.09	0.25	0.12	0.11
Access Doc. Info	52.29	23.41	0.17	2.23	0.03*	0.25	0.16	0.15
Trustworthiness Discrimination	0.37	0.40	0.07	0.92	0.36	0.13	0.07	0.06
Avg. Usefulness	-1.15	0.76	-0.14	-1.53	0.13	0.09	-0.11	-0.10
Avg. Interest	2.11	0.68	0.28	3.09	0.002**	0.23	0.22	0.21

DV: Word Count, $R^2=0.14$, Adj. $R^2=0.12$

* $p<.05$, ** $p<.01$, *** $p<.001$

The model predicting the number of arguments students generated in their open-ended responses based on multiple source use behaviors was also significant, $F(6,187) = 8.13$, $p<.001$. The model explained 20.68% of the variance in number of arguments produced. Table 14 presents a summary of the model. The total time devoted to source use, average source usefulness, and average source interestingness were all significant predictors in the model. Notably, students' average rating of source usefulness was a negative predictor.

The model predicting students' SOLO scores was also significant, $F(6, 187) = 7.56$, $p<.001$, explaining 19.53% of the variance in students' open-ended response SOLO scores. The total time students devoted to source use, accessing of document information, discrimination in trustworthiness ratings, and average interestingness ratings

were significant predictors. Table 15 presented a summary of the multiple regression model predicting SOLO scores.

Table 14

Multiple Source Use Behaviors Predicting Number Of Arguments

Variable	B	Std. Error B	β^*	t	Sig.	Correlations		
						Zero	Partial	Part
Total Source	-0.11	0.18	-0.05	-0.64	0.52	0.17	-0.05	-0.04
Total Time	0.00	0.00	0.33	4.15	0.00***	0.35	0.29	0.27
Access Doc. Info	0.49	0.52	0.07	0.94	0.35	0.20	0.07	0.06
Trustworthiness Discrimination	0.01	0.01	0.11	1.35	0.18	0.14	0.10	0.09
Avg. Usefulness	-0.06	0.02	-0.30	-3.48	0.00**	-0.04	-0.25	-0.23
Avg. Interest	0.06	0.02	0.33	3.76	0.00***	0.19	0.26	0.24

DV: No. of Arguments, $R^2=0.21$, $Adj. R^2=0.18$

* $p<.05$, ** $p<.01$, *** $p<.001$

Table 15

Multiple Source Use Behaviors Predicting SOLO Scores

Variable	B	Std. Error B	β^*	t	Sig.	Correlations		
						Zero	Partial	Part
Total Source	-0.01	0.07	-0.01	-0.10	0.92	0.24	-0.01	-0.01
Total Time	0.00	0.00	0.20	2.57	0.01*	0.31	0.18	0.17
Access Doc. Info	0.38	0.20	0.14	1.96	0.05*	0.27	0.14	0.13
Trustworthiness Discrimination	0.01	0.00	0.20	2.59	0.01**	0.25	0.19	0.17
Avg. Usefulness	-0.01	0.01	-0.12	-1.34	0.18	0.08	-0.10	-0.09
Avg. Interest	0.02	0.01	0.25	2.90	0.00**	0.22	0.21	0.19

DV: SOLO scores, $R^2=0.20$, $Adj. R^2=0.17$

* $p<.05$, ** $p<.01$, *** $p<.001$

The model predicting the total number of citations students spontaneously included in their open-ended responses based on multiple source use behaviors was overall significant, $F(6, 187) = 3.93$, $p<.01$, explaining 11.2% of variance. Only the

proportion of sources for which document information was accessed was a significant predictor in the model. A multiple regression summary table is presented in Table 16.

Table 16

Multiple Source Use Behaviors Predicting Number Of Citations

Variable	B	Std. Error B	β^*	t	Sig.	Correlations		
						Zero	Partial	Part
Total Source	0.00	0.02	-0.01	-0.11	0.91	0.13	-0.01	-0.01
Total Time	0.00	0.00	0.02	0.21	0.84	0.16	0.02	0.01
Access Doc. Info	0.19	0.06	0.26	3.43	0.00***	0.30	0.24	0.24
Trustworthiness Discrimination	0.00	0.00	0.11	1.37	0.17	0.14	0.10	0.09
Avg. Usefulness	0.00	0.00	-0.01	-0.09	0.93	0.10	-0.01	-0.01
Avg. Interest	0.00	0.00	0.12	1.33	0.19	0.15	0.10	0.09

DV: No. of citations, $R^2=0.09$, $Adj. R^2=0.06$

* $p<.05$, ** $p<.01$, *** $p<.001$

Individual Difference Factors and MSU Behaviors

Research Question 3: What is the nature of the association between individual difference factors and students' manifest source use behaviors?

Prior to running Pearson's correlation the bivariate normality assumption was evaluated by ascertaining the normality of each variable in the model. The Shapiro-Wilk test for normality, considered to be a less conservative test than the Kolmogorov-Smirnov, was used to determine normality. All of the individual difference factors and multiple source use behavioral variables, with the exception of students' average usefulness ratings, indicated non-normality (i.e., $ps<0.05$). Table 17 presents results for the Shapiro-Wilk test for normality as well as skewness and kurtosis metrics that confirm these results. While there are no firm guidelines for skewness and kurtosis, often values greater than 1 or less than -1 are considered to be concerning (Osborne, 2010). Visual

representations of variable distributions (i.e., boxplots, Normality Q-Q plots) likewise confirmed normality concerns, see Appendix P.

Table 17

Normality Statistics for Predictor Variables

Variable	Shapiro-Wilk		Skewness	Kurtosis
	Statistic	Sig.	Statistic (SE)	Statistic (SE)
Prior Knowledge	0.80	0.00	0.73 (SD=0.18)	-1.01 (SD=0.35)
Attitudes	0.72	0.00	1.33 (SD=0.18)	0.68 (SD=0.35)
Source Eval Behaviors	0.98	0.04	1.94 (SD=0.17)	10.30 (SD=0.35)
Total Sources	0.81	0.00	-0.95 (SD=0.17)	0.06 (SD=0.35)
Total Time	0.87	0.00	0.15 (SD=0.18)	-1.67 (SD=0.35)
Access Doc Info	0.81	0.00	-0.25 (SD=0.18)	0.63 (SD=0.35)
Trust Discrimination	0.97	0.00	-0.66 (SD=0.18)	0.64 (SD=0.35)
Avg. Usefulness	0.99	0.29	-0.54(SD=0.18)	0.44 (SD=0.35)
Avg. Interestingness	0.97	0.00	-0.31 (SD=0.17)	-0.74 (SD=0.35)

Due to violations of normality, rather than using Pearson’s correlation coefficient, Spearman’s rank-order correlation, a non-parametric test of associating was used (Bishara & Hittner, 2012; Gauthier, 2001). Spearman’s rho may be a more powerful test for bivariate non-normal associations than Pearson’s correlation for transformed data (Fowler, 1987; Gauthier, 2001), while data transformations may present challenges for interpretation (Osborne, 2002; Tabacknick & Fidell, 2007). Indeed, Spearman’s rho may be considered to be a data transformation in-and-of itself, as all values are assigned a rank that definitionally have a uniform shape across variables (Bishara & Hittner, 2012).

Table 18 has a correlation matrix for the relation between individual difference factors and multiple source use behaviors. Students’ self-reported source evaluation behaviors were found to be associated with total number of texts used, $\rho(188)=0.16$, $p<.05$, and with the percentage of sources for which document information was accessed, $\rho(186)=0.14$, $p=0.05$. Further, students’ attitudes were found to be positively associated

with average ratings of source usefulness, $\rho(185)=0.20, p<0.01$. Finally, each of the three individual difference factors was found to be associated with students' average ratings of source interestingness. Specifically, students' ratings of source interestingness were found to be significantly associated with prior knowledge, $\rho(187)=0.14, p<.05$, attitudes, $\rho(185)=0.30, p<.001$, and reported domain general source evaluation behaviors, $\rho(186)=0.18, p<0.05$. All of these associations corresponded to small to moderate effects.

Learner Characteristics Predicting Response Quality

Research Question 4: To what extent do individual difference factors predict response quality when students complete a multiple source task?

The fourth research question asked: To what extent do individual difference factors (i.e., prior knowledge, stance, and domain general source evaluation behaviors) predict response quality when students complete a multiple source task?

Multiple regression analysis was used. Individual differences factors (i.e., prior knowledge, stance, and domain general source evaluation behaviors) were used to predict the four measures of response quality. Separate multiple regression models were run for each outcome variable (i.e., word count, number of arguments, SOLO score, and number of citations).

The regression model predicting word count based on learner characteristics was significant, $F(3, 183) = 5.17, p<.01$, with an R^2 of 0.08. Students' attitudes were a significant predictor of word count. Table 19 presents a model summary.

Table 18

Spearman's Correlation Matrix for Individual Difference Factors and Source Use Behaviors

	1	2	3	4	5	6	7	8	9
1. Knowledge	--								
2. Attitudes	0.63***	--							
3. Source Eval Behaviors	0.26**	0.29**	--						
4. Total Source	-0.03	0.07	0.16*	--					
5. Total Time	-0.02	-0.00	0.06	0.49**	--				
6. Access Document Info	0.00	0.03	0.14*	0.25**	0.41**	--			
7. Trust Discrimination	0.07	0.12	0.08	0.46**	0.22**	0.15*	--		
8. Avg. Usefulness	0.05	0.20**	0.07	0.01	0.17*	0.16*	-0.09	--	
9. Avg. Interestiness	0.14*	0.30**	0.18*	0.07	0.21**	0.11	-0.04	0.64**	--

Table 19

Learner Characteristics Predicting Word Count

Variable	B	Std. Error B	β	t	Sig.	Correlations		
						Zero	Partial	Part
Prior Knowledge	3.34	4.48	0.07	0.75	0.46	0.18	0.06	0.05
Attitudes	21.42	7.78	0.25	2.76	0.01**	0.24	0.20	0.20
Source Eval Behaviors	-15.30	8.45	-0.14	-1.81	0.07	-0.05	-0.13	-0.13

DV: Word count, $R^2=0.08$, Adj. $R^2=0.06$

* $p<.05$, ** $p<.01$, *** $p<.001$

The model predicting the number of arguments produced was also significant $F(3, 183)=4.10, p<.01$, with $R^2=0.06$. Prior knowledge was a significant predictor of number of arguments produced. Table 20 presents a multiple regression summary table.

Table 20

Learner Characteristics Predicting Arguments

Variable	B	Std. Error B	β	t	Sig.	Correlations		
						Zero	Partial	Part
Prior Knowledge	0.28	0.10	0.25	2.85	0.01**	0.25	0.21	0.20
Attitudes	-0.00	0.17	0.00	-0.00	1.00	0.15	0.00	0.00
Source Eval Behaviors	0.01	0.19	0.00	0.02	0.98	0.06	0.00	0.00

DV: Arguments, $R^2=0.06$, Adj. $R^2=0.05$

* $p<.05$, ** $p<.01$, *** $p<.001$

Learner characteristics were not significant predictors of students' SOLO scores, $F(3, 183) = 1.91, p=0.13$, nor of the number of citations included in students' compositions, $F(3, 183) = 0.30, p=0.82$.

Learner Characteristics and Multiple Source Use Behaviors Predicting Response Quality

Research Question 5: To what extent do individual difference factors and multiple source use behaviors predict response quality when students complete a multiple source task?

The final research question was: to what extent do individual difference factors (i.e., prior knowledge, stance, and domain general source evaluation behaviors) and multiple source use behaviors (i.e., number of sources selected, time on source use, accessing document information, and source ratings) predict response quality when students complete a multiple source task?

In order to answer the third research question, a two-step multiple regression was run predicting each of the four open-ended response metrics (i.e., word count, number of

arguments, SOLO score, and number of citations) based on learner characteristics (i.e., prior knowledge, stance, and domain general source evaluation behaviors), entered in Step 1, and multiple source use behaviors, (i.e., number of sources selected, time on source use, accessing document information, and source ratings) entered in Step 2. All information presented in regression summary tables is based on full models, with both learner characteristics and multiple source use behaviors entered as predictors.

The model predicting word count based on learner characteristics and multiple source use behaviors was significant, $F(9, 175) = 4.68, p < .001$, explaining 19.40% of the variance in word count. Attitudes, domain general source evaluation behaviors, accessing document information, and average interestingness ratings were all found to be significant predictors in the model. The variable, domain general source evaluation behaviors, was a negative predictor. Table 21 presents a summary of the multiple regression model. The model including both learner characteristics and multiple source use behaviors ($R^2=0.19$), offered a significant improvement in fit over predicting word count via learner characteristics alone ($R^2=0.07$), $\Delta R^2=0.12, \Delta F(6,175) = 4.46, p < 0.001$.

Likewise, the model for number of arguments produced was significant, $F(9, 175) = 5.71, p < .001$, explaining 22.71% of the variance in number of arguments included in students' responses. Prior knowledge, time on texts, and average ratings of source usefulness, and source interestingness were all significant predictors in the model. Source usefulness was negatively associated with number of arguments produced. The joint model ($R^2=0.23$), offered a significant improvement in fit over predicting arguments via learner characteristics alone ($R^2=0.06$), $\Delta R^2=0.17, \Delta F(6,175) = 6.29, p < 0.001$. Table 22 has a model summary of the full multiple regression model.

Table 21

Learner Characteristics and Multiple Source Use Behaviors Predicting Word Count

Variable	B	Std. Error B	β	t	Sig.	Correlations		
						Zero	Partial	Part
<i>Step 1: Learner Characteristics</i>								
Prior Knowledge	3.76	4.19	0.08	0.89	0.37	0.17	0.07	0.06
Attitudes	16.67	7.48	0.20	2.23	0.03*	0.23	0.17	0.15
Source Eval Behaviors	-21.52	7.95	-0.20	-2.71	0.01**	-0.06	-0.20	-0.18
<i>Step 2: Multiple Source Use Behaviors</i>								
Total Source	3.55	7.97	0.04	0.45	0.66	0.17	0.03	0.03
Total Time	0.00	0.00	0.12	1.52	0.13	0.22	0.11	0.10
Access Doc. Info	56.27	24.07	0.18	2.34	0.02*	0.23	0.17	0.16
Trustworthiness Discrimination	0.27	0.41	0.05	0.67	0.51	0.14	0.05	0.05
Avg. Usefulness	-1.15	0.77	-0.14	-1.49	0.14	0.10	-0.11	-0.10
Avg. Interest	1.85	0.72	0.24	2.57	0.01*	0.23	0.19	0.18

DV: Word count, $R^2=0.19$, $Adj. R^2=0.15$; * $p<.05$, ** $p<.01$, *** $p<.001$

Table 22

Learner Characteristics and Multiple Source Use Behaviors Predicting Number Of Arguments

Variable	B	Std. Error B	β	t	Sig.	Correlations		
						Zero	Partial	Part
<i>Step 1: Learner Characteristics</i>								
Prior Knowledge	0.27	0.09	0.24	2.90	0.00**	0.25	0.21	0.19
Attitudes	-0.02	0.17	-0.01	-0.15	0.88	0.14	-0.01	-0.01
Source Eval Behaviors	-0.14	0.18	-0.06	-0.79	0.43	0.05	-0.06	-0.05
<i>Step 2: Multiple Source Use Behaviors</i>								
Total Source	-0.04	0.18	-0.02	-0.21	0.84	0.16	-0.02	-0.01
Total Time	0.00	0.00	0.28	3.52	0.00**	0.31	0.26	0.23
Access Doc. Info	0.51	0.54	0.07	0.96	0.34	0.18	0.07	0.06
Trustworthiness Discrimination	0.01	0.01	0.08	0.99	0.32	0.14	0.08	0.07
Avg. Usefulness	-0.05	0.02	-0.29	-3.16	0.00**	-0.02	-0.23	-0.21
Avg. Interest	0.05	0.02	0.30	3.26	0.00**	0.20	0.24	0.22

DV: No. of Arguments, $R^2=0.23$, $Adj. R^2=0.19$; * $p<.05$, ** $p<.01$, *** $p<.001$

The full modeling predicting students' open-ended SOLO scores based on learner characteristics and multiple source use behaviors was significant, $F(9, 175) = 5.33$, $p < .001$, explaining 21.53% of variance. While individual difference factors were not significant predictors in the model, total time devoted to source use, discrimination in trustworthiness ratings, and average ratings of interestingness were all significant predictors. The model including both learner characteristics and multiple source use behaviors ($R^2=0.22$), offered a significant improvement in fit over predicting word count via learner characteristics alone ($R^2=0.03$), $\Delta R^2=0.19$, $\Delta F(6, 175)=6.95$, $p < 0.001$. The multiple regression summary is presented in Table 23.

Table 23

Learner Characteristics and Multiple Source Use Behaviors Predicting SOLO Scores

Variable	B	Std. Error B	β	t	Sig.	Correlations		
						Zero	Partial	Part
<i>Step 1: Learner Characteristics</i>								
Prior Knowledge	0.03	0.04	0.08	0.90	0.37	0.13	0.07	0.06
Attitudes	0.06	0.06	0.09	1.01	0.32	0.15	0.08	0.07
Source Eval Behaviors	-0.12	0.07	-0.12	-1.70	0.09	-0.01	-0.13	-0.11
<i>Step 2: Multiple Source Use Behaviors</i>								
Total Source	0.02	0.07	0.02	0.23	0.82	0.24	0.02	0.02
Total Time	0.00	0.00	0.21	2.60	0.01*	0.31	0.19	0.17
Access Doc. Info	0.37	0.21	0.14	1.79	0.08	0.25	0.13	0.12
Trustworthiness Discrimination	0.01	0.00	0.18	2.30	0.02*	0.25	0.17	0.15
Avg. Usefulness	-0.01	0.01	-0.12	-1.32	0.19	0.09	-0.10	-0.09
Avg. Interest	0.02	0.01	0.24	2.53	0.01*	0.22	0.19	0.17

DV: SOLO scores, $R^2=0.22$, *Adj. R*²=0.18

* $p < .05$, ** $p < .01$, *** $p < .001$

A model predicting the total number of citations in students' responses based on learner characteristics and multiple source use behaviors was significant, $F(9, 175) = 2.43$, $p < 0.05$, with an $R^2=0.11$. The only significant predictor in the model was the

proportion of document information accessed. A model summary is presented in Table 24.

Table 24

Learner Characteristics and Multiple Source Use Behaviors Predicting Number of Citations

Variable	B	Std. Error B	β	t	Sig.	Correlations		
						Zero	Partial	Part
<i>Step 1: Learner Characteristics</i>								
Prior Knowledge	0.01	0.01	0.06	0.68	0.50	0.06	0.05	0.05
Attitudes	0.00	0.02	-0.01	-0.15	0.88	0.03	-0.01	-0.01
Source Eval Behaviors	-0.02	0.02	-0.09	-1.17	0.24	-0.02	-0.09	-0.08
<i>Step 2: Multiple Source Use Behaviors</i>								
Total Source	0.00	0.02	0.01	0.05	0.96	0.14	0.00	0.00
Total Time	0.00	0.00	0.04	0.44	0.66	0.17	0.03	0.03
Access Doc. Info	0.18	0.06	0.24	2.98	0.00**	0.28	0.22	0.21
Trustworthiness Discrimination	0.00	0.00	0.11	1.34	0.18	0.16	0.10	0.10
Avg. Usefulness	0.00	0.00	-0.02	-0.22	0.82	0.09	-0.02	-0.02
Avg. Interest	0.00	0.00	0.12	1.26	0.21	0.13	0.09	0.09

DV: Citations, $R^2=0.11$, Adj. $R^2=0.07$

* $p<.05$, ** $p<.01$, *** $p<.001$

CHAPTER 5: DISCUSSION

Multiple source use or the selection, processing, and evaluation of multiple information sources to meet task goals (Gil et al., 2010b) has been identified as a crucial competency for today's students (Goldman & Scardamalia, 2013a) and correspondingly, has become an increasingly prominent part of undergraduate curricula (Metzger et al., 2003). At the same time, undergraduate students have been found to have limited skills with regard to multiple source use (Grimes & Boening, 2001) perhaps because of how few students report receiving formal training in online source use and evaluation (Burton & Chadwick, 2000).

Theoretical models of multiple source use may provide insights into understanding the challenges that learners may experience with text engagement. Such models have conceptualized the multiple source use process as unfolding through a series of iterative steps involving source selection, source processing, and source evaluation (Rouet & Britt, 2011). While each of these steps has been previously examined independently in the empirical literature (e.g., sources selection: Kammerer et al., 2009; source evaluation: Bråten et al., 2009), limited work has considered a joint model involving all three of these processes impacting open-ended response quality. Further, while learner characteristics (i.e., individual cognitive resources that students bring with them to a MSU task, Rouet & Britt, 2011) have been theorized as important to students' subsequent text engagement, these have been examined only to a limited extent.

The purpose of the present study was to address these gaps in the literature. First, this study sought to employ multiple source use behavioral indicators in a comprehensive model predicting open-ended response quality. Second, this study examined the effects

of individual difference factors on open-ended response quality, both by themselves and in conjunction with multiple source use behaviors. Further, this study examined associations between individual difference factors and multiple source use behaviors. Individual difference factors and multiple source use behaviors were examined in the context of students completing an open-ended task about a contemporary political issue using both traditional (e.g., newspapers) and more novel sources (e.g., blog, Twitter, Wikipedia).

This chapter begins with a discussion and interpretation of findings associated with each of five research questions empirically investigated. Next synthesis of research findings and general conclusions will be presented, particularly as they inform understandings of multiple source use as put forth in the MD-TRACE model (Rouet, 2006) and the conceptual model guiding this study, presented in Chapter 1. Finally, study limitations and future directions are discussed.

Conclusions and Implications

Research Question 1

The first research question was concerned with describing students' multiple source use process and product. Overall, students seemed fairly engaged or effortful in their source use, accessing the majority of sources for substantial periods of time. As expected, students rated sources high in reliability (e.g., analysis essay, public opinion survey) as most trustworthy. Conversely, twitter and the blog post were rated as significantly less reliable, despite providing on-the-ground information and first-hand witness accounts about the Arab Spring in Egypt. In history tasks, students have generally found primary sources to be more reliable than secondary accounts (Stahl et al., 1996). However, ratings may have been impacted by new, digital source types having a

reputation of being low in reliability (Lim, 2009). Further, the twitter account and the blog were clearly biased: the twitter account was the official English language account of the Muslim Brotherhood, while the blog strongly favored el-Sisi. At the same time, the analysis essay, rated high in trustworthiness, was comparably biased against General el-Sisi, perhaps due to the author who is regarded as an authoritative source. This demonstrates the complexity of trustworthiness ratings as contingent on students' judgments of source type, bias, authors' perspective, and expertise. Results seem to suggest that students evaluate texts primarily based on considerations of source type, consistent with prior research suggesting that students rely on source type-based schema in forming trustworthiness evaluations (Bråten et al., 2011). The newspaper article, intended to be a source moderate in reliability, was rated most trustworthy. This was unexpected as newspapers may present ambiguous cues about reliability to readers (Armstrong & Collins, 2009). Also, in prior work newspapers have been rated lower in trustworthiness than more authoritative texts written by experts (Strømsø et al., 2011). However, the topic of the task may have led students to consider the newspaper to be a discipline-specific and therefore particularly trustworthy source. This may indicate that in addition to using source-type based heuristics in judging sources (Brem et al., 2001) students may also use discipline-specific heuristics in evaluating texts. Strømsø et al. (2011) suggest that newspapers may be considered trustworthy because they are easily comprehensible, enabling students to evaluate the credibility of their content.

Trustworthiness ratings seemed to be associated with students' source selections. The newspaper, rated as the most trustworthy source, was also the text selected most often. Conversely, the blog post was visited least often and considered to be a source low

in trustworthiness. This correspondence provides initial support for the hypothesis that students generate heuristic, source type-based trustworthiness judgments prior to text selection and that these then serve as the filter through which texts are processed (Brem et al., 2001). Students' heuristic-based source evaluations are concerning given the limited frequency with which students accessed document information. Indeed, students considered author and publication information for less than half of sources accessed. Moreover, students were less likely to access document information for digital source types (i.e., Wikipedia, twitter), which may have particularly ambiguous authorship (Britt & Gabrys, 2002) and therefore require more thorough evaluation (Coiro, 2003b). Students rated these sources as lowest in trustworthiness, which they appeared to be.

As in prior research (Rouet et al., 1997) students' evaluations of source trustworthiness largely corresponded to their ratings of source usefulness as well as to the amount of time they devoted to source use. There were two deviations from this pattern. Students found Wikipedia to be a useful source, spending considerable time on it, despite rating it as one of the less trustworthy texts. Lim (2009) found similar results, explaining that students' familiarity with and positive feelings toward Wikipedia were responsible for its use, despite students not considering it to be a particularly trustworthy source. Conversely, despite rating the public opinion poll as quite trustworthy, students did not consider it to be particularly useful and did not devote much time to its use. This may be because the public opinion poll contained quantitative information, which has been found to be difficult for students to process (Shah, Mayer, & Hegarty, 1999). More generally, this may suggest, first, that, indeed, usefulness, like relevance, is defined not only in relation to task but also by appropriateness to learners' skill level (McCrudden & Schraw,

2007). Second, usefulness may be the deciding factor in the amount of time learners devote to source use. Curiously, interest was not found to differ by source type. This may suggest that interest was functioning at a more general level of topic or task.

In examining the association between source access and response quality, across source types, accessing texts was associated with greater response elaboration (i.e., number of arguments produced). This is perhaps because with each additional source accessed, students gained more information to incorporate into their responses.

Accessing the newspaper article and the public opinion survey, in particular, served not only to increase the number of arguments produced but also students' SOLO scores or response integrativeness and evaluativeness. The newspaper article was a relatively balanced source of information, providing critiques of both Morsi and el-Sisi. This may have encouraged greater response integration and the evaluation of both sides of the issue. Meanwhile the public opinion poll provided seemingly objective data that may have helped students in generating more evaluative responses when comparing Morsi and el-Sisi.

Overall, students seemed to be generating quality responses. Measures of word count and the number of arguments produced suggest that students' reasoning was elaborated and that learners were providing a number of justifications in support of their positions. Further, the majority of SOLO scores fell into the three-to-four range, indicating that students were integrating information from multiple sources and evaluating conflicting information across texts. However, spontaneous sourcing was limited, with almost half of students not referencing the origin of information included in

their responses at all. This outcome reinforces concerns about students' use of document information (Britt & Aglinskias, 2002; Perfetti et al., 1999).

In examining the frequency with which various source types were cited, digital source types, rated as lower in trustworthiness (e.g., blog, Wikipedia, twitter), were cited less often, while students cited sources considered to be more trustworthy (e.g., newspaper, analysis essay) to a greater extent. Differences in citation rates across sources were not uniformly proportionate to the number of students using texts or accessing document information. This may indicate that while students do use sources they rate as low in trustworthiness in the research process and in composing their responses, they resist formally citing such texts. Head and Eisenberg (2010), in a survey of undergraduates' Wikipedia use, similarly found that although undergraduates commonly use Wikipedia, they rarely cite it in their academic work due to professors' cautions against using Wikipedia as a source.

Research Question 2

The second research question examined the extent to which multiple source use behaviors were predictive of task performance. Multiple regression analyses determined that a number of source use behavioral metrics were predictive of response quality measures. All four regression models based on multiple source use behaviors were significant, explaining between 9.19% of variance in the number of spontaneous citations included to 20.68% of variance in the number of arguments generated. Accessing document information and average text interestingness ratings were predictive of word count. Total time on texts, and students' average ratings of source usefulness and source interestingness, were associated with the number of arguments produced. Finally SOLO

scores, considered to be the open-ended response metrics most reflective of high-level engagement in multiple source use (i.e., with points assigned for information integration and the evaluation of conflicting positions) were found to correspond to a variety of multiple source use behavioral indicators. Specifically, total time on texts, accessing document information, discrimination in trustworthiness ratings, and average interestingness ratings all predicted SOLO scores. In particular, the two metrics associated with an epistemic orientation toward source use (i.e., accessing document information and trustworthiness discrimination) were associated with learners' SOLO scores. Accessing document information was the only significant predictor of number of citations produced. Table 25 presents a summary of significant predictors in each model.

Students' average rating of source interestingness was predictive of all open-ended response quality metrics, with the exception of number of citations included. This variable was a significant predictor of elaborative response metrics, word count, and number of arguments produced, as well as of SOLO scores. This outcome underlines the importance of interest not only in source engagement but also in task outcome (Bråten & Strømsø, 2006).

Evaluations of source trustworthiness and usefulness seemed to impact students' responses in somewhat different ways. Accessing document information and discrimination in trustworthiness ratings were both associated with SOLO scores, corresponding to the extent to which students' written products reflected multiple text integration and information evaluation. Usefulness ratings were associated with the number of arguments produced or the elaborativeness of students' responses. This suggests that while task-directed source evaluations (i.e., usefulness) may be involved in

generating responses, more epistemic, or text-directed evaluations (i.e., trustworthiness) are needed to truly engage in the higher level cognitive processes associated with multiple source use (Britt et al., 1999; Rouet, 2006). In particular, as suggested by the Documents Model of Multiple Source Use, accessing document information was associated with greater integration and evaluation of conflicting information (Perfetti et al., 1999), as reflected in students' SOLO scores.

In particular, attending to author and publication information during source use (i.e., accessing document information) is believed to support students' linking of information across sources in their responses while evaluating competing claims put forward across texts (Britt et al., 1999). Likewise, better distinguishing between sources high and low in reliability (i.e., discrimination in trustworthiness ratings) resulted in higher SOLO scores, potentially by promoting the corroboration of conflicting information. Trustworthiness discrimination scores indicated that, at the least, students were sensitive to differences between texts; this may have resulted in their comparison of sources offering competing evidence. In other words, texts may be useful when they help students aggregate enough information to formulate a sufficiently elaborated open-ended response. However, to integrate and evaluate information, students must consider document information and further evaluate trustworthiness across texts.

While not a significant predictor of SOLO scores, the average rating of source usefulness, was a negative predictor of the number of arguments generated. This finding seemed counterintuitive. It was expected that students would determine usefulness ratings based on their perceptions of sources providing information closely connected to task demands (Rouet & Britt, 2011). Students rating sources as more useful may have

based such judgments on being able to locate task-relevant arguments in texts with ease, potentially including a greater number of arguments in their responses as a result. In fact, the opposite effect was found. One possible explanation for the negative association between usefulness ratings and number of arguments generated comes from McNamara et al.'s (1996) work. Specifically, McNamara et al. (1996) found that for high-knowledge students, texts requiring inferencing or deeper processing were better comprehended than easier texts. Likewise, in the present study, if students found sources to be less useful (i.e., providing less task-relevant information) they may have been driven to think more deeply about the information in texts or to draw on prior knowledge to a greater extent. Despite finding sources less useful in explicitly providing task-relevant arguments, these students may have, nonetheless, generated a greater number of arguments in their responses. While McNamara et al. (1996) found only high knowledge students to benefit from less comprehensible texts, less useful sources were more widely beneficial in our sample. It may be the case that students in our sample were sufficiently experienced with multiple source tasks to capitalize on implicit arguments presented in texts, even when they considered sources to be of limited usefulness (i.e., providing few explicit task-relevant arguments).

Time on texts was a significant predictor of both the number of arguments produced and students' SOLO scores. However, this metric may be interpreted in a number of ways. Time on texts may parallel interest as a measure of engagement or may simply reflect the volume of information students were able to gather from sources accessed. Alternately, time on texts has been termed a measure of effort expenditure

(Bråten et al., 2014) and may correspond to the cognitive effort students applied during source use and depth of processing.

Despite a hypothesized association, the number of sources selected was not a significant predictor of any of the four outcome measures in this study. There are at least three possible explanations for this missing association. First, it may be the case that number of sources accessed is not an effective summative score of students' source selection. More individualized metrics of whether or not students accessed each specific text may need to be used. Alternately, since the majority of students in this study accessed the majority of sources (i.e., more than 4, on average), this predictor may have had insufficient variability. Given findings that much of situation model construction happens once students access two sources, with more limited model developmental occurring after accessing of a third source (Stahl et al., 1996), it may be that the specific number of texts students accessed after the third, be it four, five, or six, was not as important. Finally, as all of the sources available to students in the library were selected to be relevant (i.e., to provide information that could be used in formulating a response), it may be that source selection was not a driver of response quality in this study. Indeed, it has been hypothesized that source selection matters most in ascertaining relevance (Anmarkrud et al., 2013; Rouet & Britt, 2011).

Table 25

Summary of Predictors

Predictors	Word Count	Arguments	SOLO	Citations
<i>Learner Characteristics</i>				
Prior Knowledge		Model 2 Model 3		
Attitudes	Model 2 Model 3			
Source Evaluation Behaviors	Model 3			
<i>Multiple Source Use Behaviors</i>				
Total Sources				
Total Time		Model 1 Model 3	Model 1 Model 3	
Access Document Information	Model 1 Model 3		Model 1	Model 1
Trustworthiness Discrimination			Model 1 Model 3	
Avg. Usefulness		Model 1 Model 3		
Avg. Interest	Model 1 Model 3	Model 1 Model 3	Model 1 Model 3	

Note: Model 1: Multiple source use behaviors predicting response quality; Model 2: Learner characteristics predicting multiple source use behaviors; Model 3: Learner characteristics and multiple source use behaviors predicting response quality

Research Question 3

The third research question examined the associations between individual difference characteristics and students' multiple source use behaviors. In general, relations were modest. As expected, students' domain general source evaluation behaviors had a small association with the total number of sources students accessed as

well as with the percentage of sources for which document information was accessed.

This was expected as these two behavioral indicators mapped on to scale items.

Specifically, students' reports of the frequency with which they corroborated sources may have manifested in their accessing more texts. Items asking students about the frequency with which they verified author credentials may have manifest in the percentage of sources for which students accessed document information. However, reported source evaluation behaviors were not associated with discrimination in trustworthiness ratings, suggesting the limitations in correspondence between reported and demonstrated source evaluations (Hargittai et al., 2010).

Prior knowledge and attitudes were not associated with many of the source use behavioral indicators (e.g., total number of sources accessed, total time on texts).

However, all three individual difference factors were found to be associated with average interestingness. As the students in our sample were generally limited in prior knowledge and did not hold particularly strong attitudes toward the Arab Spring, it is reasonable to suppose that the examined individual difference factors were not sufficiently pronounced such as to manifest in multiple source use behaviors. Rather, it seems that prior knowledge and stance only served to promote general task engagement, as expressed in ratings of source interestingness.

The relation between attitudes and source usefulness may reflect the ease which students were able to find information relevant to either position (i.e., in favor of Morsi or el-Sisi). The sources included in the library offered a variety of perspectives on the target prompt, such that students holding stronger attitudes with regard to the Arab Spring in

Egypt would have been able to easily access attitude consistent information. This type of motivated source use has been documented in prior research (Nickerson, 1998).

Research Question 4

The fourth research question examined the extent to which individual difference factors were associated with performance on the open-ended multiple source use task. However, only limited relations were identified. Stance was associated with word count, whereas prior knowledge was associated with the number of arguments students produced. Models predicting SOLO scores and the number of citations included in students' responses were not significant. Both of the outcome measures that individual difference factors predicted were elaborative metrics (i.e., word count and number of arguments produced). This may suggest that prior knowledge and stance support information aggregation but not necessarily higher-level processing, as would have been reflected in students' SOLO scores. More precisely, it may be the case that the students in our sample did not have a sufficient level of knowledge or attitudinal strength to support deeper level processing, such as information integration. In fact, Britt et al. (1999) have suggested that document model construction, or the integration and reconciliation of information across multiple texts, is not the predominant comprehension approach favored by students. Typically, learners are more likely to simply engage in information gathering. Nonetheless, many students' responses did reflect information integration and an evaluative approach but this may have stemmed from source use behaviors rather than from internal cognitive resources.

Alternately, there may be a measurement limitation. Prior knowledge and attitudes have been said to function as schema or filters for students' interactions with

texts. The measures used to assess prior knowledge and attitudes in this study may not have been sufficiently complex to capture students' schematic understanding or organizational structure for information. For instance, the identification of terms prior knowledge assessment can be said to best assess learners' familiarity with the topic rather than a comprehensive knowledge framework, *per se*.

Research Question 5

Research question five investigated a full model, including individual difference factors and multiple source use behaviors, as predictive of the open-ended response metrics. This paralleled models examined for research questions two and four. In this case, individual difference factors can be considered to be control variables. Controlling for individual difference factors (i.e., prior knowledge, stance, and domain general source evaluation behaviors), multiple source use behaviors were nonetheless predictive of open-ended response performance. Additionally, examining partial correlations for predictors in each of the full models allowed for the identification of factors driving variance explained in each outcome metric.

In predicting word count, in the full model, students' reported domain general source evaluation behaviors had a negative relation with word count. Further, examining partial correlations, determined that students' reported domain general source evaluation behaviors were the strongest predictor in the model. This relation is difficult to interpret, in part because it was not a significant predictor when only learner characteristics were included in the model. One possibility is that the bivariate relation between students' domain general source evaluation behaviors and average ratings of source usefulness, a negative but not significant predictor of word count, was affecting the role of source

evaluations behaviors as a predictor in the model. Alternately, it may be the case that students who are more disposed toward expending cognitive effort on source evaluation may not have had sufficient cognitive bandwidth left over to generate elaborative responses. However, source evaluation behaviors were not negatively associated with the other outcome variables, lessening the plausibility of this explanation. Another explanation may be that students who were more concerned with source evaluation were also less verbose. Certainly such a conclusion is supported by the semi-partial correlation between source evaluation behaviors and word count, removing the effects of other factors from both.

In the model predicting of number of arguments generated, total time on texts and average ratings of source interestingness were the strongest predictors in the model. Among the individual difference factors, only prior knowledge was a significant predictor. All three of these variables may have facilitated the volume of information students were able to gather, resulting in a greater number of arguments produced. For instance, students having more time to spend on sources and being more immersed in texts, as reflected by average ratings of source interestingness, may have simply been more able and motivated to accumulate information toward formulating a response.

SOLO scores were predicted only by multiple source use behaviors, specifically, time on texts, trustworthiness discrimination, and students' average ratings of source interestingness. As with the number of arguments produced, time on texts and interestingness were the two strongest predictors in the model. This likely reflects scores zero to two on the SOLO taxonomy being contingent on the number of reasons or arguments students produce in favor of their position. Then, students' discrimination in

trustworthiness ratings and proportion of sources for which document information was accessed were the strongest predictors in the model. Except for time on texts, these two sourcing-related metrics had the highest zero-order correlations with SOLO scores. The importance of these variables' association with SOLO scores may reflect the SOLO taxonomy assigning higher scores for integrative responses and those including information evaluation.

As with other models, students' average rating of source interestingness was a significant predictor of SOLO scores, while prior knowledge and attitudes were not significant in the model. A prior study examining individual difference factors as part of a model of multiple source use found somewhat different results. Specifically, Bråten et al. (2014) examined both the direct and indirect effects of prior knowledge and situational interest on multiple source use. They found prior knowledge to have a direct effect on multiple text comprehension but not on behavioral variables (i.e., time on texts and situational interest). Situational interest had only an indirect effect on multiple text comprehension via students' reported deep strategy use. More work is needed to disentangle the direct and indirect effects of individual difference factors on multiple text engagement and comprehension.

Only accessing document information was associated with the number of citations students included in their open-ended responses. This relation seems straightforward in that students concerned with evaluating document information when engaged in source use then also referenced this information when writing. Pragmatically, accessing document information (e.g., author) provided students with the information they needed to generate citations. Of course, it is possible that students were only accessing document

information for citation purposes, to meet writing conventions, rather than to evaluate sources.

In addition to examining each research question individually, conclusions may be drawn across analyses, particularly with regard to the extent to which findings from this study align with the MD-TRACE model, the guiding framework for this study.

MD-TRACE Model

The MD-TRACE model is a behaviorally focused conceptualization of multiple source use that also acknowledges the role of learner characteristics in impacting source engagement. Based on the models examined in this study, there seems to be initial support for a behavioral focus on multiple source use. In fact, in models jointly considering learner characteristics and behavioral factors as predictive of performance, it was source use behaviors that dominated. The importance of source use behaviors was particularly evident in models predicting SOLO scores that evaluated students' responses based on their integrativeness and evaluativeness.

The MD-TRACE model focuses on three specific sub-steps (i.e., source selection, processing, and evaluation in the service of integration) that define students' interactions with texts (Rouet & Britt, 2011). In current analyses, there was support for the effects of source processing and evaluation on task performance, with the role of source selection being more tenuous. It may be the case that source selection is less important when all texts available to students are task-relevant and therefore that source selection functions as a screening for relevance (Rouet & Britt, 2011). Time on texts proved to be both an efficient metric of processing and a strong predictor of a variety of outcome measures. Students' source evaluation was of particular interest in this study and all three of the

evaluation dimensions students were asked to consider (i.e., trustworthiness, usefulness, and interestingness) proved to be differently but significantly associated with response quality. However, there is remaining concern regarding the extent to which students readily engage in source evaluation, given the limited percentage of sources for which students accessed document information.

Across models, students' average ratings of source interestingness was the most consistent predictor, significantly contributing to variance explained in word count, number of arguments produced, and SOLO scores. At the same time, interest ratings did not differ across source type, suggesting that interest ratings may have functioned as more general measures of task or topic engagement. This is consistent with prior research (Bråten & Strømsø, 2006; Wigfield & Guthrie, 2000), which has highlighted the importance of interest and engagement to performance on complex tasks. Findings from the present study certainly point to the possible merit of including interest as a component in the MD-TRACE model. More work is needed to determine whether interest is best examined as a cognitive resource promoting students' performance or as a situational factor, arising from task affordances. In MD-TRACE terms, more research is needed to ascertain whether interest constitutes a transitory or permanent cognitive resource (Rouet & Britt, 2011).

Wade, Buxton, and Kelly (1999) have used the term, *reader-text interest*, to signify learners' motivation for text engagement. They identify this kind of interest as arising from learners' *individual interest* and *cognitive interest*. Individual interest may be considered to be learners' long-term, personal, disposition toward a topic (Hidi, Renninger, & Krapp, 2004; Schiefele, 1991). Cognitive interest has been defined as the

amount of learning a passage induces, with learning considered to be a motivating experience in-and-of itself (Kintsch, 1980). Both forms of interest may need to be further incorporated into models of multiple source use.

Most surprising in our analyses was the limited role of individual difference factors in task performance. These findings may be explained in a number of ways through the lens of the MD-TRACE model. First, it may be the case that the MD-TRACE model is best considered to be a representation of novice source use. When prior knowledge and attitudes are limited, task performance may indeed be more behaviorally driven. It may only be with greater expertise that individual difference factors shape the nature of text engagement. For instance, prior work has found disciplinary experts to be distinct in their source evaluation (Wineburg, 1991) and Rouet et al. (1997) have found source use to differ between graduate students expert in a target domain and graduate students who were experts but in a different discipline.

It may also be the case that, while individual difference factors do affect the multiple source use process, the quality of learners' written responses is behaviorally determined. Certainly the MD-TRACE model suggests that students' written products are developed primarily in reference to task demands, rather than based on background factors. This explanation is partially supported by students' post-hoc ratings of task engagement being associated with both individual difference factors and learners' text evaluations, but not with response quality metrics.

Additionally, it may be that learner characteristics need to be assessed in more task-specific ways. Just as in the present study interest was assessed as specific to texts, rather than as general topic interest, it may be that prior knowledge and attitudes likewise

need to be assessed in more text-focused ways or as directly embedded in students' text processing. In this study, students were placed in a multiple text context and instructed to use sources in formulating their responses (even though not all of them elected to do so). Perhaps either these instructors or the volume of information in sources, as compared to their more limited knowledge and attitudinal valence, caused behavioral factors to have a more pronounced effect.

Conceptual Model

The conceptual model, introduced in Chapter 1 as the guiding framework for this study, posits that individual difference factors (i.e., prior knowledge, stance on the target issue, and domain general source evaluation behaviors) and multiple source use behaviors (i.e., source selection, source processing, and source evaluation) jointly predict multiple text task performance. There seems to be at least partial support for the proposed conceptual model.

Considering the role of learner characteristics in task performance determined that learners' prior knowledge, stance, and domain general source evaluation behaviors were all significant predictors. In particular, stance, or students' strength of commitment to a position with regard to the target issue was related to the number of words produced. Students holding stronger feelings for or a greater commitment toward a specific position may have produced more elaborated or expressive responses; however, these may not necessarily have been of higher quality. Strength of stance was neither associated with the number of arguments produced nor with response integrativeness or evaluativeness, as reflected in SOLO scores. Such a response profile may have resulted from students with strong stances on the target issue engaging in case-building, or development of one

sided argumentation accompanied by a dismissal of competing evidence (Alexander et al., 1998; Nickerson, 1998). For instance, one student's response was:

My family is Egyptian and this is an issue that hits close to home. I believe that General El Sisi should remain in power till the country is back in order...In the newspaper article, it stated that General El Sisi helped capture 11 Muslim Brotherhood members because they were trying to 'incited violence against police.' Although the US response was of terror that the Egyptian people are running out of their civil rights, I believe we take just as much precaution as they do. Recently, a girl on twitter tweeted that she was part of a terrorist group and will do something big in June, America Airlines forwarded her name to the FBI and she was going to be investigated because she posed a threat to our national security. This is a similar situation. The Muslim Brotherhood is a violent extremist terrorist group who is willing to do anything to get their point across. The possible hint that they might hurt a police officer should be enough to arrest them; especially since "social media has become the main form of communication" for the terrorist group. General El Sisi is trying his best to keep Egypt safe. It is better to be safe than sorry. The Egyptian government ensures that they will be given a fair trial, just like they would in the U.S. General Sisi is trying his best to take out the Muslim Brotherhood because a terrorist group should not be running a country which is something Morsi could not do.

Although the student's generated response was quite lengthy and elaborated, only a limited number of unique arguments were presented and evidence was evaluated only to a limited extent, typical of case-building (Alexander et al., 1998; Nickerson, 1998). The response this student produced is consistent with her initial stance, reported prior to task engagement, in favor of General el-Sisi.

As expected (Alexander et al., 1994; Wolfe & Goldman, 2005), prior knowledge was found to be associated with the number of arguments students produced. This likely reflects students with greater stores of prior knowledge being better able to process multiple texts (Symons & Pressley, 1993). Reduced cognitive effort during text processing may have resulted in students being able to accumulate a greater volume of information and therefore generate a larger number of arguments in their responses.

However, prior knowledge did not have an impact on SOLO scores. This lack of an association may stem from the sample being overall limited in prior knowledge, or from other factors, such as relational reasoning or writing ability, being implicated in students' multiple source integration and response composition. In other words, prior knowledge may be key to the amount information students draw from texts but other individual difference factors, requiring further investigation, may impact how learners elect to integrate and reproduce information.

Unlike the relations proposed in the conceptual model, individual difference factors did not seem to function in tandem or to universally impact performance metrics. Rather it appears that various aspects of learners' backgrounds had an impact on different facets of performance. The finding that particular individual factors are differentially associated with various dimensions of task performance may explain why students' competence in multiple source use has been found to be limited, even at the undergraduate level (Grimes & Boening, 2001; Metzger et al., 2003). In fact, MSU's categorization as more cognitively effortful and complex than single text processing (Bråten & Strømsø, 2006), may reflect that it requires students to coordinate prior knowledge, topic stance, domain general source evaluation behaviors, and other individual difference factors in order to complete tasks.

Three types of learner characteristics were examined in the present study. While students' prior knowledge and topic stance may be considered to be cognitive and motivational factors, reported domain general source evaluation practices constituted a behavioral metric. The extent to which these different types of background factors (i.e., cognitive, motivational, and experiential) ought be considered as predictive of task

performance remains an open question. While prior knowledge and topic stance were positively associated with performance, students' reported engagement in domain general source evaluation behaviors was a negative predictor of word count. At the same time, correlation analyses determined that source evaluation behaviors were positively associated with high-level source use behaviors as manifest in students accessing more sources, potentially to corroborate information, and considering document information to a greater extent. Such inconsistent findings further emphasize the importance of examining both process (i.e., multiple source use behaviors) and product (i.e., task performance). Additionally, multidimensional aspects of students' responses need to be considered to understand which learner characteristics may make a difference for varied facets of task performance.

Across performance metrics, students' average rating of text interestingness was both the most consistent and the strongest predictor. As text interestingness has not been investigated in prior research on multiple source use, this is a novel and promising finding. There was a positive association between text interestingness and word count, number of arguments produced, and SOLO scores, indicating that the motivational benefits attributed to text-based interest, indeed, resulted in improved performance (Hidi, 1990; Hidi & Baird, 1986). Average ratings of text-based interest were significantly correlated with another multiple source use indicator, time on texts, predictive of the number of arguments students generated and SOLO scores. The association between text interestingness and time on texts has been established in prior research on single text processing (e.g., Kelly & Belkin, 2001; Morita & Shinoda, 1994) and provides insight into the role of interest in multiple source use. Interest seems to support more persistent

text engagement, resulting in students aggregating more information from texts and being able to produce more extensively justified responses as a consequence. Further, ratings of interestingness predicting SOLO scores provides support for Hidi's (2000) contention that text-based interest not only allows for sustained engagement but also for a deeper level of processing. Put another way, evaluations of text interestingness and time on texts seemed to function in concert supporting text engagement. This engagement can be thought to include both motivational absorption and cognitive intentness, resulting in deeper processing and higher SOLO scores. Where Bråten et al. (2014) have considered time on text to be a measure of effort expended, this study evidences that such effort is accompanied by and predicated on interest.

Work by Rouet et al. (1996), as well as the proposed conceptual model, suggests that students' evaluations of source trustworthiness and usefulness correspond to one-another. As a contrast, in this study, ratings of source trustworthiness and source usefulness seemed to function independently and relate differently to varied performance outcomes examined. Discrimination in trustworthiness ratings was a positive predictor of SOLO scores, while average rating of source usefulness was a negative predictor of the number of arguments generated. In fact, this divergence in evaluations may reflect students approaching the given task with two different goal orientations. Students more discriminating in trustworthiness ratings and accessing document information to a greater extent may have been seeking to compose responses that integrated information across sources and evaluated evidence presented in texts. Indeed, these two measures (i.e., trustworthiness discrimination and accessing document information) were positive predictors of students' SOLO scores. Students who found sources to be low in usefulness

may have been looking to identify a single “answer” explicitly presented in text. These students may have been dissatisfied by the lack of a response directly stated in the sources available and, as a result, may have rated texts accessed as low in usefulness or as limited in aiding them in meeting their task goals. Students who found texts to be of low usefulness may have elected to abandon source use and develop responses based on prior knowledge. For instance, one student constructed the response:

The United States of America should support General El Sisi for several reasons having to do with halting the spread of Islamic radicalism. If Morsi were to take over as president of Egypt this would pose as a huge problem for the U.S. because Morsi is an Islamic extremist who would aid Al-Qaeda in military operations and pose as a very serious threat to the national security of our closest ally in the region, Israel. If Morsi were to become president then Al-Qaeda would just have another place to hide from U.S. forces and Egypt has a powerful military, which could provide valuable military assets to Al-Qaeda. Also, this would lead to the spread of Islamic extremists because the schools in Egypt would be forced to teach an intolerant and radical form of Islam. As I mentioned before Israel's security would also be at stake because imagine if Egypt attacked Israel and Israel responds, this creates an insurmountable amount of unrest in the already unpredictable Middle East, which could lead to oil prices sky rocketing, which can cripple our already fragile and recovering economy. Overall it would just be disastrous for everyone if Morsi were to be reinstated as the president of Egypt.

Although a number of arguments are presented in this answer, this response was not based on information offered in the source library. Specifically, none of the texts included information about the implications of the Arab Spring for Israeli security. This student may have found the sources in the library to be limited in usefulness and, therefore, have generated a number of arguments based on prior knowledge. Such a response also demonstrates limited integration of conflicting information or evaluation of evidence, suggesting that students generating arguments independent of library sources

and findings texts to be of limited usefulness may have been, at least partially, reluctant to engage in high-level source use.

From a statistical standpoint, it may be the case that trustworthiness discrimination scores and average ratings of source usefulness capture different properties of multiple source evaluation or that averaging ratings of source usefulness diluted the precision of this factor.

There seemed to be greater concordance in students' multiple source use behaviors impacting response quality than was found among the individual difference factors. Text interestingness and time on texts, an indicator of source processing, worked together to predict a variety of performance metrics. However, as far as the evaluative measures were concerned, accessing of document information, trustworthiness discrimination scores, average ratings of source usefulness, and source interestingness were associated only to a limited extent with one-another and were non-coherent in predicting various performance measures. This suggests that, in fact, these evaluative dimensions may be tapping quite different aspects of multiple source use experience and that greater exactness is required to understand *which* evaluative processes matter for *what*.

As has been suggested throughout the conclusions section, there are a number of limitations in the present study that prompt future work.

Limitations and Future Directions

Limitations discussed address study design, assessment methods, generalizability of findings, and factors to consider in future research.

Behavioral Indicators

A primary strength of this study, its focus on behavioral variables, is also a limitation. This study introduced novel source use metrics to track students' text engagement within an online interface (e.g., whether or not students elected to access document information). Although such metrics proved to be an efficient and minimally obtrusive way to capture source processing in a large sample of students, they also present challenges in interpretation. For instance, time on text may be interpreted as a measure of effort expended (Bråten et al., 2014), depth of processing (Rouet et al., 2011), or engagement. There is much to be gained from more qualitative analyses of multiple source use to compliment behavioral measures. Think alouds, in particular, may provide additional validity and insight into the interpretation of behavioral metrics and have been shown to be effective methods for capturing and classifying students' strategic processing (Mason et al., 2010b; Pressley & Afflerbach, 1995).

Further, summative measures of source engagement (e.g., average usefulness ratings) may mask the variability of individuals' interactions with texts. In analyses, each source students visited contributed to overall metrics computed, even if learners did not find a text useful or elected to disregard information from a particular source in response composition. More qualitative analyses may be better able to access and disentangle specific reader-text interactions within a broader multiple text context. Further, the conceptual model proposed as well as multiple regression as an analysis framework are predictive in nature. While manifest relations were effectively modeled, the explanatory power of these is limited. More qualitative work may support better explaining the

processes of multiple source use, their interaction, and their results in product than does predictive modeling.

The multiple source use literature originates, in part, from theories of intertextuality, or the notion that students' experiences with one source are intrinsically related and impacted by their interactions with all previously encountered texts (Shuart-Faris & Bloome, 2004). Although summative measures of multiple source use do treat learners' experiences across texts as a single body or one task experience, qualitative methods may be preferable for tracing the development and revision of students' situation models across sources. In particular, intertextual perspectives may have particular bearing on data, such as those collected as part of this study, that are gathered throughout students' multiple text task engagement rather than at a single point-in-time (e.g., source evaluations following task completion, Bråten et al., 2009; Bråten et al., 2011). As source use metrics were collected as students used each source they accessed, there is reason to believe that texts accessed earlier and students' judgments of such texts impacted subsequent source selections, processing, and evaluations. This is consistent with the MD-TRACE model that lays out source processing behaviors (i.e., source selection, source processing, and source evaluation) as iterative and in interaction with one another.

Different perspectives on the consistency of students' source use have been offered. Morita and Shinoda (1994) have found students' source access to be fairly uniform, both in terms of times spent on texts and interestingness ratings. As a contrast, Rouet and Britt (2011) suggest that particularly in a multiple text context, students may liberally alternate between texts varying in usefulness and trustworthiness, until

identifying desired sources. Little work has been done to empirically investigate the consistency of students' multiple source use and the ways that students' interactions with texts may be impacted by their previous experiences (Shuart-Faris & Bloome, 2004), both within a task and with texts more generally. This type of work may require more qualitative analysis or fine-grained examination of students' dynamic interactions with texts and source access pathways. While such analyses were not conducted in the present study, data collected can support such investigation in the future.

Examining Strategies

The present study was limited in its consideration of students' source use strategies that indeed, may have an effect on both process and product (Bråten et al., 2014; Strømsø et al., 2003). In the multiple source use literature, methods to assess strategic processing have been limited, focusing on self-report questionnaires (Bråten & Strømsø, 2011) or the interpretation of navigation data under restrictive task conditions (Reader & Payne, 2007). In part, this may stem from difficulties in identifying the nature of strategies specific to multiple source use or source use on the Internet (Afflerbach & Cho, 2009). Think alouds or retrospective interviews may be initial fruitful directions for identifying the breadth of strategies students may employ when interacting with multiple texts (Walraven et al., 2009).

Increasingly, there has been an interest in understanding the self-regulatory and metacognitive monitoring strategies that students use during multiple source use (Bråten & Strømsø, 2005). As with strategies more generally, more needs to be understood about which specific self-regulatory and metacognitive techniques are most effective for multiple source task completion. While software and systems have been developed to

prompt students' self-regulation and metacognition during multiple source use (e.g., AutoTutor, Azevedo, 2005; Met.A.Ware, Stadtler & Bromme, 2007, 2008), these may benefit from a greater theoretical grounding. In the present study, while self-regulation and metacognition were not directly assessed, prompting students to evaluate each source after access may have served as a metacognitive intervention stimulating more reflective source use and may have changed students' interactions with subsequent texts. As participants in this study constituted a non-expert sample, they may have been particularly susceptible to this type of evaluative prompting. Future work can adopt a two-pronged approach, using think alouds and eye-tracking methods to better understand the strategies students employ during source use and manipulating interfaces to improve students' text engagement.

Optimal Non-Expert Multiple Source Use

An additional challenge in defining what may constitute strategic multiple source use is that there are limitations in developmental understandings of this process. While there are expert models of multiple source use (Rouet et al., 1997; Wineburg, 1991), research has been more limited in defining optimal, non-expert source use (Bråten et al., 2011). One possibility for identifying the type of text engagement that may be preferable for K-12 or post-secondary students is using task performance to identify low-level and high-level learners, or those students benefitting least and most from multiple source use, and then examining these contrasting groups for differences in processing (e.g., Goldman et al., 2012). Data collected as a part of this study may be extended in a similar way. For one, it is possible to identify students with responses especially high and low in quality and examine differences in processing. Alternately, a post-task prior knowledge measure

can be administered to students to examine differences in “learning,” or knowledge gains from pre- to post-task.

Initial studies have, indeed, examined differences in pre- and post- task assessments of topic knowledge as measures of learning (Braasch & Goldman, 2010). For instance, Goldman et al. (2012) administered a 20 item true-false knowledge of volcanic eruptions assessment to students prior to and following their engagement in a multiple text task, with those students exhibiting the greatest and least change in performance selected for further investigation. However, during task completion, in addition to gains in topic knowledge, students also likely develop experience with and learn how to engage in MSU. Students’ learning of multiple source use processes and strategies through task completion has received more limited attention in the literature. While Britt and Aglinskias (2002) and Sanchez, Wiley, and Goldman (2012) have examined the extent to which students’ sourcing practices transfer from one multiple text task to another, these evaluative behaviors were compared following intervention. Further work is needed to understand how, through the course of multiple text task engagement, even absent intervention, students develop skills and strategies and learn how to engage in multiple source use.

The focus on growth in topic knowledge as a consequence of multiple source use is understandable. MSU has been described as a task driven process where, often, performance is contingent on gaining knowledge about an issue (e.g., Bråten et al., 2009). From an assessment standpoint, more methods are needed for determining not only whether students satisfy task demands but also whether learners thoughtfully and deliberately engaged with texts. For instance, Grimes and Boening (2001) suggest that

instructors grade written products, paying limited attention to the processes students used in acquiring information and to students' selections of credible sources. Given students' high levels of confidence (Grimes & Boening, 2001) and manifest limitations (Britt & Aglinskias, 2002; Metzger et al., 2003) in their multiple source use skills, more formative assessments are needed. Instructor evaluations of students' multiple source use process and greater encouragement of learner reflection may promote ultimate performance. Further, more process focused and incremental assessment may inform developmental perspectives on multiple source use. Assessment models should further seek to capture the complexity and domain and task nested nature of multiple source use. Specifically, there is a need to identify effective multiple source use practices that carry across tasks and domains as well as to isolate discipline specific approaches to information use (e.g., Wiley et al., 2009; Wineburg, 1991).

In addition to developing viable models of competent, non-expert source use, the extent to which students at various levels of development are capable of the cognitive processes associated with multiple source use remains an open question. This study focused on undergraduate students, who, despite experiencing challenges with multiple source use, have been considered to be sufficiently cognitive mature to engage in the deliberative and evaluative processes that high-level multiple source use demands. At the same time, studies of multiple source use have been conducted with children as young as 6th grade (e.g., Wallace & Kupperman, 1997; Wallace, Kupperman, Krajcik, & Soloway, 2000), although their multiple source use processing may be best characterized as information location. For instance, Wallace et al. (2000) described 6th grader's source use as motivated by goals including *finding the perfect page* that would give them all of

the information needed and *finding an answer* rather than engaging in the integrative and evaluative processes associated with deliberative multiple source use. However, the majority of studies have focused on undergraduate students, with even high school learners found to be markedly more limited in exhibiting high-level multiple source use (Britt & Aglinskias, 2002; Wineburg et al., 1996). Nonetheless, more work is needed to better define the level of source use that may be expected from younger students and to develop pedagogies that may foster earlier reasoning with multiple texts.

Expanded Analyses

The analyses used in this study and factors examined could be expanded. First, only direct effects were investigated in this study. Given prior work modeling both direct and indirect effects in multiple source use (Bråten et al., 2014) and the limited direct effects found between individual difference factors and task product, examining indirect effects is a logical next step. Individual difference factors may affect general task engagement, manifesting as students' ratings of source interestingness, which then affects performance.

The individual difference factors assessed in this study need to be further examined. The prior knowledge measure used can be modified to better capture students' schematic understanding of the topic. For instance, asking students to describe the Arab Spring in Egypt in a more open-ended fashion may better assess both the volume of what students know as well as how coherent and organized their knowledge is. In particular, assessment of students' stance can be improved by using a scale to increase reliability, beyond a single item. An interesting area for future work is to determine the extent to which engagement in a multiple source use task served to change students' stance with

regard to the topic. This is a promising direction for research since findings from this study determined that following multiple source use students were better able to commit to a position (i.e., in favor of Morsi or el-Sisi) than they were prior to task engagement. The extent to which this change reflects a change in attitudes or simply a change in knowledge has yet to be determined.

A number of additional individual difference factors ought be likewise included in future models. The first of these is students' domain and topic interest, particularly given the importance of in situ text interestingness ratings identified in current analyses. Likewise reading level or source comprehensibility ratings would help inform the extent to which fluidity in processing at the text-base level is facilitative of higher level multiple source use. A number of researchers have advocated examining single-text reading abilities (e.g., fluency) as a predictor of students' abilities to engage with multiple texts without undue cognitive load or disorientation (Britt et al., 1999; Strømsø et al., 2008). Additionally, epistemic beliefs, examined widely in the multiple text literature, should be incorporated into future models (Bråten, Britt, Strømsø, & Rouet, 2011; Ferguson, Bråten, & Strømsø, 2012). In particular, Bråten et al. (2014) have introduced an epistemic belief dimension specific to multiple source use, students' beliefs in the *justification of knowledge by multiple texts*, that merits further investigation.

Design Parameters

In constructing the multiple source task, a number of factors were delimited. Students did not self-identify texts. Rather they were restricted to accessing only six, researcher-selected sources in the library. Future studies should examine students' search as a component of the multiple source use process. Learners were also prompted to

evaluate each source accessed along researcher-specified dimensions. The extent to which students may spontaneously evaluate sources remains unclear, as do the effects of prompting students to evaluate texts. Based on prior work, it can be expected that such prompting increased students' source evaluations, overall, (Gerjets et al., 2011) potentially improving performance and increased the frequency with which document information was accessed. Students were asked to rate sources according to specific dimensions, specified by the research (e.g., trustworthiness, usefulness, interestingness). It remains unclear the extent to which participants' folk definitions of these constructs aligned with those drawn from the literature. For instance, theoretically, students' trustworthiness evaluations are considered to be predicated on students' sourcing and evaluation of document information (Britt et al., 1999; Perfetti et al., 1999). However, in this study, as prompted to, students evaluated the trustworthiness of each source used, despite not accessing document information for half of the texts selected.

This suggests that trustworthiness judgments may take two forms: they may be deliberative and based on available document information (Britt & Aglinskis, 2002) or they may be more heuristic, conveying students' general perceptions of reliability (Brem et al., 2001). Heuristic-based source judgments may have been dependent on source type. Source type was the only piece of source information available to students absent their accessing additional document information (Bråten et al., 2009). Heuristic judgments may have also been based on genre assumptions. In this study, unlike in prior research, students most frequently accessed the newspaper sources, rated it highest in trustworthiness, and proportionately, cited it to the greatest extent in their responses. At the same time, students did not access document information about the newspaper source

at a higher rate. This suggests that a genre-based heuristic may have affected students' source use and evaluation. The newspaper may have been the source most frequently accessed and rated as most trustworthy because the task addressed a contemporary political issue

Further, both source integration and information evaluation were assessed only in the extent to which they were reflected in students' responses. While it is typical to assess multiple source processing via a performance measure (e.g., Britt & Sommer, 2004; Gil et al., 2010), there nonetheless may be a disconnection between students' cognitive text processing or situational model construction and demonstrated response composition. Undergraduates have, indeed, been found to sometimes struggle with academic writing (Flowerdew, 2000; Grimes & Boening, 2001) and in the present study, there was no measure of writing skills.

Generalizability

There are questions with regard to how much findings from the current study can be generalized beyond the current sample as well as beyond the current task. First, the sample was not fully representative of the University student population in terms of gender, minority status, and class standing. Second, differences in multiple source use were found across these demographic categories. While examining demographic factors was beyond the scope of the present study, it is certainly an exciting area for future research. For instance, while the role of gender in multiple source use has been examined only to a limited extent, conflicting findings have been found. While Cooper and Weaver (2003) found males to be more comfortable in digital contexts, Lycke, Strømsø, and Grøttum (2002) did not find gender differences in male and females' online source use.

More recently, Bråten and Strømsø (2006) found females to report more multiple text processing strategies during source use. Explanatory frameworks for these gender differences have been limited, although prior research in traditional print contexts has found females to be more linguistically facile (e.g. Riding & Al-Sanabani, 1998).

In addition to examining gender, given recent findings of disparities in technology access for minority and non-minority students (Attewell, 2001), this is an important area for further examination, particularly as it manifests at the undergraduate level.

Disparities in minority and White students' performance in online contexts may be exacerbated as minority students not only have more limited access to technology but have also been found to be more limited in demonstrating the reading comprehension skills necessary for effective multiple source use (Leu, Kinzer, Coiro, & Cammack, 2004). Moje, Ciechanowski, Kramer, Ellis, Carrillo, and Collazo (2004) have suggested that minority students may experience challenges with disciplinary literacy because traditional classroom texts fail to reflect diverse students' knowledge and out-of-classroom experiences. The disconnect between academic texts and minority students backgrounds may be manifest in this study by minority students, overall, considering sources to be of lower trustworthiness than did their White peers. At the same time, the online environment may facilitate minority student multiple source use. Digital contexts may serve as a *third space* allowing minority students to connect disciplinary content (*secondary space*) with *first space* personal experience gained outside of the classroom (Bhabha, 1994; Moje et al., 2004; Soja, 1996).

The MD-TRACE model suggests that students' multiple source use is impacted not only by internal cognitive resources but also by external resources, providing a

context for the task. Students in this study were effortful in their multiple source use, producing quality responses. There are a variety of external resources that may have facilitated this type of task engagement. The study was completed in a computer laboratory, with students at individual stations, working on the task simultaneously. This may have contributed to making task completion feel like it was unfolding within the context of a class or traditional academic setting, rather than as part of a laboratory experiment. Further, the researcher was present and stated the directions for participants, in addition to their being presented in the study interface. This may have contributed to students' adherence to instructions asking them to provide evidence to justify their responses. Given the length of time students devoted to source use, texts selected may have been particularly appropriate for use with a low knowledge undergraduate sample in terms of length, difficulty, and volume and variability in information presented. However, the extent to which students may perform differently at home, in a classroom setting, or on a graded assignment remains unclear. Finally, separating the research and response phases may have resulted in students being more systematic in their source use. This manipulation may have led participants to focus on gathering sufficient information to compose a well-reasoned response, rather than on simply generating an answer. In some cases, students were paid for participation. Different compensation methods for the sample (e.g., extra credit, payment) may have differently affected students' motivation for study completion.

The task was about a contemporary event, prominent in the news. This task topic proved sufficiently interesting to involve students in extended multiple source use resulting in quality responses and moderately high post-task ratings of engagement.

Specific insights into multiple source use when responding to a contemporary topic were gained. These include students' preference for topic-specific source types (i.e., newspapers) and distrust of digital source types (e.g., Twitter, Wikipedia), even when these may be quite pertinent. Further, students were found to be especially limited in accessing document information to evaluate new, online sources (i.e., Twitter, Wikipedia), despite these having ambiguous authorship, and to be hesitant in citing such texts in their responses. Digital source types may feature particularly prominently in contemporary topics; therefore, these findings may be especially relevant to understanding how source use in response to current events may unfold. Nevertheless, as the extent to which students access document information for various source types has not been examined in prior research (e.g., Bråten et al., 2009; Wiley et al., 2009), these findings contribute to understandings of students' multiple source use outside of contemporary political issues as well. Given evidence that multiple source use has a strong domain specific component (Rouet et al., 1997; Wiley et al., 2009; Wineburg, 1991) processing may be different for other topics and domains, requiring further investigation.

There remain questions regarding the frequency with which multiple source tasks are present in the classroom. Hicks, Doolittle, and Lee (2012) point out that only a limited body of work has examined teachers' selections of textual materials for their classrooms. In the domain of history, there has been practitioner-focused advocacy for teachers to incorporate more and multiple primary sources into the curriculum (Hartzler-Miller, 2001; Wineburg, 2001; VanSledright, 2002). Yet, these initiatives have impacted the history classroom only to a limited extent. In a survey of high school history

teachers, Hicks et al. (2012) found that only 64.6% of teachers reported their students “comparing and contrasting details across multiple data sources” *often* or *very often*. Further, only 48.6% of teachers reported their students “interrogating historical data given the context of the data’s creation” *often* or *very often*. In fact, teachers themselves report difficulties with multiple source use or, more specifically, with identifying effective online sources to introduce into the classroom. Specifically, 36.7% of teachers *agreed* or *strongly agreed* with the statement: “it is frustrating locating useful sources on the web.” Further, 58.3% of teachers endorsed there being “too many web sites to locate suitable primary sources” as an *important* or *very important* reason for why they did not use web-based primary history sources.

The extent to which multiple sources are incorporated into classrooms in domains beyond history remains a question. A survey from the Pew Research Center (Purcell, Heaps, Buchanan, & Friedrich, 2013) offers some insights. History and social studies teachers were the most likely, of secondary teachers surveyed to, “look for material online to create lesson plans.” Of history teachers, 46% reported doing so, as compared to 39% of science teachers, 33% of English and Language Arts teachers, and 23% of math teachers. This suggests that history teachers may have greater facility with incorporating multiple sources, or at the least, sources found online, into the classroom as compared to teachers in other domains. Beyond considering the use of multiple texts in the classroom, more work is needed to document the type of multiple text tasks teachers, across domains, present to students and the frequency with which such tasks are introduced.

Implications for Practice

The importance of effective multiple source use has already been recognized in contemporary curricula and testing. For instance, the English Language Arts Common Core standards for 11th-12th grade ask that students be able to, “Integrate and evaluate multiple sources of information presented in different media or formats (e.g., visually or quantitatively) as well as in words in order to address a question or solve a problem” (Common Core State Standards, 2010, CSS.ELA-Liteacy.RI.11-12.7). However, teaching in light of such standards remains a challenge. This study suggests at least three ways that teachers may be helped in meeting Common Core demands.

First, there is a need to incorporate non-traditional source types into the classroom and to teach students how to use and evaluate these effectively. Second, there is reason to suggest that reading for longer improves task performance. Often, in real-world multiple text contexts, students rapidly sequence through sources in an effort to find *the answer* or a single source that can meet all of their information needs (Wallace et al., 2000; Walraven et al., 2009). Encouraging students to read, not only scan, sources and to engage with texts for longer may result in higher level processing and performance.

Finally, text interest was found to be quite important across models of multiple source use, meriting further empirical examination and more focused inclusion in models of multiple source use. However, there are encouraging findings for un-interesting source use as well. Many times adults are required to read texts they are not interested in, about which they have limited knowledge, and do not hold strong beliefs. This happens when reading about important events in the news, selecting a health insurance plan, or signing up for a credit card.

An encouraging finding from this study is that diligent source use behavior (i.e., accessing document information, discrimination between high and low trustworthiness sources) may be enough to support multiple text integration and information evaluation. Behavioral source use factors were found to result in successful task performance even though our sample was limited in knowledge and strength of stance. This suggests that teaching students how to use texts effectively may be sufficient to produce desired source engagement and performance. Programs providing undergraduates with a source evaluation “checklist” (Calkins & Kelley, 2007) or training students to attend to author information (Britt & Aglinskias, 2002) hold promise for implementation in the classroom.

In this study, multiple source use proved to be a highly effective exercise. Students with limited knowledge and ill-defined stance, after engaging with multiple texts for a limited period of time (i.e., 15 minutes on average), were able to produce reasoned, elaborated, integrated, open-ended responses about a complex issue. Multiple source use has long been heralded as offering benefits over single text comprehension (Wiley & Voss, 1999), so perhaps the greatest instructional implication this study offers is evidence of the merits of increasing the preponderance of multiple texts in the classroom.

APPENDICES

Appendix A: Demographic Characteristics and Multiple Source Use

Appendix A includes differences by gender, minority status, and class standing in learner characteristics, multiple source use, and performance, as well as predictive models of open-ended response quality, controlling for these demographic factors.

Table 26

Differences in Learner Characteristics, Multiple Source Use, and Performance by Gender

	Males (n=59)	Females (n=129)
<i>Learner Characteristics</i>		
Knowledge	2.71 (2.65)	2.16 (2.59)
Attitudes	2.22 (1.71)	1.98 (1.45)
Domain General Source Evaluation Behaviors	4.44 (1.01)	4.49 (1.25)
<i>Multiple Source Use Behaviors</i>		
Total Number of Sources	4.22 (1.70)**	4.94 (1.36)**
Total Time on Texts	12.17 min. (7.89)***	17.18 min. (9.38)***
Access Doc. Information	0.34 (0.43)*	0.49 (0.40)*
Trist Discrimination	51.16 (26.99)	56.43 (26.05)
Avg. Usefulness	57.54 (14.42)	62.01 (16.19)
Avg. Interestingness	58.09 (15.18)	59.77 (17.93)
<i>Performance</i>		
Word Count	181.39 (152.92)	193.80 (126.02)
Arguments	4.64 (3.44)	4.84 (2.84)
SOLO Scores	2.53 (1.09)	2.86 (1.15)
Citations	1.36 (2.19)	1.78 (1.81)
<i>Post-Task Engagement</i>		
Engagement	4.27 (1.61)	4.29 (1.66)

* $p < .05$, ** $p < .01$, *** $p < .001$

Females used significant more sources than males, $t(\text{adj. } 93.01) = 2.85$, $p < 0.01$, (Males: $M = 4.22$, $SD = 1.70$; Females: $M = 4.94$ $SD = 1.36$) and devoted significantly more total time to text engagement, $t(186) = 3.57$, $p < 0.001$, (Males: $M = 12.17$ minutes, $SD = 7.89$; Females: $M = 17.18$ minutes, $SD = 9.38$). Further, females accessed document information in a greater percentage of cases, $t(186) = 3.57$, $p < 0.001$, (Males: $M = 0.34$, $SD = 0.43$; Females: $M = 0.49$, $SD = 0.40$).

Table 27

Differences in Multiple Source Use and Performance by Minority Status

	White (n=97)	Minority (n=91)
<i>Learner Characteristics</i>		
Knowledge	2.47 (2.67)	2.19 (2.57)
Attitudes	2.12 (1.55)	1.98 (1.53)
Domain General Source Evaluation Behaviors	4.34 (1.11)	4.63 (1.24)
<i>Multiple Source Use Behaviors</i>		
Total Number of Sources	4.75 (1.39)	4.67 (1.63)
Total Time on Texts	15.40min (7.40)	15.83 min (10.87)
Access Doc. Information	0.43 (0.42)	0.45 (0.41)
Trust Discrimination	58.50 (24.75)*	50.81 (27.62)*
Avg. Usefulness	61.39 (14.69)	59.84 (16.89)
Avg. Interestingness	61.76 (15.79)*	56.58 (18.12)*
<i>Performance</i>		
Word Count	204.58 (125.90)	174.26 (142.61)
Arguments	5.11 (2.66)	4.42 (3.37)
SOLO Scores	3.03 (0.99)**	2.46 (1.22)**
Citations	2.03 (2.10)**	1.24 (1.67)**
<i>Post-Task Engagement</i>		
Engagement	4.24 (1.56)	4.33 (1.72)

* $p < .05$, ** $p < .01$, *** $p < .001$

As compared to White students ($M=58.50$, $SD=24.75$), minority students had significantly lower trustworthiness discrimination scores ($M=50.81$, $SD=27.62$), $t(186)=2.01$, $p < 0.05$. Further, minority students ($M=56.58$, $SD=18.12$) rated source as significantly lower in terms of interestingness than did White students ($M=61.76$, $SD=15.79$), $t(184)=2.08$, $p < 0.05$. In terms of performance, minority students ($M=2.46$, $SD=1.22$) had significantly lower SOLO scores as compared to White students ($M=3.03$, $SD=0.99$), $t(\text{adj. } 173.50)=3.50$, $p < 0.01$. Minority students also included significantly fewer citations in their responses ($M=2.03$, $SD=2.10$) than did White students ($M=2.03$, $SD=2.10$), $t(\text{adj. } 173.50)=3.50$, $p < 0.01$.

Table 28

Differences in Multiple Source Use and Performance by Class Standing

	Freshman	Sophomore	Junior	Senior	Special Students
<i>Learner Characteristics</i>					
Knowledge	2.24 (2.51) ^a	1.91 (2.53) ^b	2.52 (2.71) ^c	1.95 (2.32) ^d	5.10 (2.73) ^{a,b,c,d}
Attitudes	2.27 (1.57)	1.84 (1.38)	2.09 (1.64)	2.09 (1.61)	2.00 (1.33)
Source Evaluation Behaviors	4.69 (1.02)	4.29 (1.22)	4.39 (1.25)	4.54 (1.27)	4.69 (0.82)
<i>Multiple Source Use Behaviors</i>					
Total Source	4.69 (1.30)	5.12 (1.18)	4.39 (1.77)	4.97 (1.46)	3.80 (1.69)
Total Time	18.51 min (9.80) ^a	17.25 min (7.33)	13.95 min (6.63)	14.76 min (12.45)	8.61 min (6.19) ^a
Access Doc. Information	0.55 (0.44)	0.44 (0.41)	0.41 (0.37)	0.43 (0.43)	0.31 (0.45)
Trustworthiness Discrimination	58.12 (24.13)	61.00 (21.97)	50.44 (27.93)	54.97 (28.69)	36.60 (27.92)
Avg. Usefulness	63.34 (17.89)	61.13 (14.67)	61.46 (15.20)	58.55 (15.71)	51.08 (11.72)
Avg. Interestingness	60.40 (18.96)	61.87 (12.37)	58.00 (18.99)	58.08 (17.79)	54.29 (14.40)
<i>Performance</i>					
Word Count	238.86 (128.72) ^a	184.74 (115.46)	173.30 (140.24)	200.26 (144.43) ^b	55.80 (47.47) ^{a,b}
Arguments	5.74 (2.99) ^a	5.09 (3.32)	4.0 (2.65)	4.90 (2.78)	2.60 (3.53) ^a
SOLO Scores	3.13 (0.86) ^{a,b}	2.85 (1.01) ^c	2.48 (1.23) ^a	2.94 (1.15) ^d	1.55 (1.23) ^{b,c,d}
Citation	1.76 (1.76)	1.40 (1.56)	1.26 (1.67)	1.87 (2.05)	0.40 (0.67)
<i>Post Task Engagement</i>					
Engagement	4.31 (1.54)	4.34 (1.72)	4.31 (1.78)	4.10 (1.61)	4.50 (1.27)

One way analysis of variance determine that students' class standing was a significant differentiator of students' knowledge, $F(4, 183) = 3.56, p < 0.01$. Post hoc comparisons using Bonferroni's method, determined that special undergraduate students had significantly higher levels of prior knowledge ($M = 5.10, SD = 2.73$) than did freshman ($M = 2.24, SD = 2.51$), $p < 0.05$, sophomores ($M = 1.91, SD = 2.53$), $p < 0.01$, juniors ($M = 2.52, SD = 2.71$), $p < 0.05$, and seniors ($M = 1.95, SD = 2.32$), $p < 0.01$. Further, students with different levels of class standing significantly differed in terms of number of sources accessed, $F(4, 183) = 2.70, p < 0.05$, and total time devoted to source use, $F(4, 183) = 3.53, p < 0.01$. Bonferroni's post hoc comparisons determined that freshman ($M = 18.51, SD = 9.80$) spent significantly more time on source use than did special students ($M = 8.60, SD = 6.19$), $p < 0.05$. Word count, $F(4, 183) = 4.45, p < 0.01$, number of arguments, $F(4, 183) = 3.31, p < 0.05$, and SOLO scores, $F(4, 183) = 5.52, p < 0.001$, also significantly differed by class standing. Freshman ($M = 238.86, SD = 128.72$), $p < 0.01$, and seniors' ($M = 200.26, SD = 144.43$), $p < 0.05$, responses had a significantly higher word count than did those of special undergraduate students ($M = 55.80, SD = 47.47$). Freshman ($M = 5.74, SD = 2.99$) also produced significantly more arguments ($M = 2.60, SD = 3.53$) than did special students, $p < 0.05$. Finally, freshman ($M = 3.13, SD = 0.86$) had higher SOLO scores than juniors ($M = 2.48, SD = 1.22$), $p < 0.05$, and special students ($M = 2.76, SD = 1.14$), $p < 0.01$; sophomores ($M = 2.85, SD = 1.01$), $p < 0.01$, and seniors ($M = 2.94, SD = 1.15$), $p < 0.01$, likewise had significantly higher SOLO scores than special students ($M = 1.55, SD = 1.23$).

Table 29

Differences in Multiple Source Use and Performance by Area of Study

	Natural Sciences	Social Sciences	Humanities	Undeclared
<i>Learner Characteristics</i>				
Knowledge	2.41 (2.78)	2.38 (2.56)	2.60 (3.06)	1.46 (2.12)
Attitudes	1.96 (1.54)	2.18 (1.61)	1.70 (1.25)	1.58 (0.90)
Source Evaluation	4.37 (1.21) ^a	4.47 (1.16) ^b	5.51 (0.72) ^{a,b}	4.20 (1.20)
<i>Behaviors</i>				
<i>Multiple Source Use Behaviors</i>				
Total Source	5.00 (1.43)	4.63 (1.53)	4.50 (1.65)	4.46 (1.45)
Total Time	15.82(11.70)	14.91(7.34) ^a	13.17(7.43)	22.55(11.09) ^a
Access Doc. Information	0.48 (0.40)	0.40 (0.40)	0.60 (0.47)	0.54 (0.49)
Trustworthiness	54.80 (27.86)	55.77 (25.90)	47.00 (27.83)	52.46 (24.62)
Discrimination				
Avg. Usefulness	58.03 (16.05)	61.71 (14.55)	63.78 (19.82)	59.26 (21.00)
Avg. Interestingness	58.17 (16.66)	60.01 (16.47)	60.33 (22.50)	55.99 (20.95)
<i>Performance</i>				
Word Count	179.19 (139.44)	191.57 (132.82)	192.40 (162.07)	218.31 (119.30)
Arguments	5.07 (3.33)	4.66 (2.84)	3.20 (2.35)	5.77 (3.63)
SOLO Scores	2.77 (1.14)	2.74 (1.13)	2.40 (1.51)	3.08 (1.00)
Citation	1.20 (1.62)	1.61 (1.86)	2.00 (1.63)	1.08 (1.12)
<i>Task Engagement</i>				
Engagement	4.31 (1.60)	4.34 (1.67)	3.90 (1.42)	4.04 (1.78)

One way analysis of variance determined that students majoring in various subject areas differed significantly in reported domain general source evaluation behaviors, $F(3, 183)=3.01, p<0.05$. Post hoc comparisons using Bonferroni's method determined that students in the humanities reported engaging in source evaluation behaviors significantly more frequently ($M=5.51, SD=0.72$) than did students in the natural sciences ($M=4.37, SD=1.21$) and the social sciences ($M=4.47, SD=1.16$). Further students majoring in various subject areas differed with regard to time spent on text use, $F(3,184)=3.00, p<0.05$. Post-hoc analyses determined that students who were undeclared devoted significantly more time to text use ($M=22.55, SD=11.09$) than did students in the social sciences ($M=14.91, SD=7.34$).

The model using demographic factors, learner characteristics, and multiple source use behaviors in predicting word count was significant, $F(12, 172) = 4.16, p < 0.001, R^2 = 0.23$. A model summary is presented in Table 30.

Table 30

Demographics, Learner Characteristics, and Multiple Source Use Behaviors Predicting Word Count

Variable	B	Std. Error B	β	t	Sig.
<i>Step 1: Demographic Factors</i>					
Gender (1=Female)	9.64	20.12	.03	.48	0.63
Minority Status (1=Non-White)	-19.07	18.37	-0.07	-1.04	0.30
Class Standing	-17.72	7.88	-0.16	-2.25	0.03*
<i>Step 2: Learner Characteristics</i>					
Prior Knowledge	5.40	4.22	0.11	1.28	.20
Attitudes	14.90	7.44	0.18	2.00	.047*
Source Eval Behaviors	-20.17	7.98	-0.18	-2.53	.01*
<i>Step 3: Multiple Source Use Behaviors</i>					
Total Source	6.07	8.00	0.07	0.76	0.45
Total Time	0.00	0.00	0.08	0.97	0.34
Access Doc. Info	55.73	23.85	0.18	2.337	0.02*
Trustworthiness Discrimination	0.03	0.41	0.01	0.07	0.95
Avg. Usefulness	-1.34	0.77	-0.16	-1.73	0.09
Avg. Interest	1.77	0.72	0.23	2.45	0.02*

DV: Word Count, $R^2 = 0.23, Adj. R^2 = 0.17$

* $p < .05$, ** $p < .01$, *** $p < .001$

The model using demographic factors, learner characteristics, and multiple source use behaviors in predicting number of arguments produced was significant, $F(12, 172) = 4.94, p < 0.001, R^2 = 0.26$. A model summary is presented in Table 31.

Table 31

Demographics, Learner Characteristics, and Multiple Source Use Behaviors Predicting Number of Arguments

Variable	B	Std. Error B	β	t	Sig.
<i>Step 1: Demographic Factors</i>					
Gender (1=Female)	0.08	0.45	0.01	0.17	0.87
Minority Status (1=Non-White)	-0.49	0.41	-0.08	-1.19	0.24
Class Standing	-0.38	0.18	-0.15	-2.18	0.03*
<i>Step 2: Learner Characteristics</i>					
Prior Knowledge	0.30	0.09	0.27	3.23	0.00***
Attitudes	-0.06	0.17	-0.03	-0.39	0.70
Source Eval Behaviors	-0.10	0.18	-0.04	-0.58	0.56
<i>Step 3: Multiple Source Use Behaviors</i>					
Total Source	0.02	0.18	0.01	0.13	0.90
Total Time	0.00	0.00	0.24	2.98	0.00***
Access Doc. Info	0.51	0.53	0.07	0.96	0.34
Trustworthiness Discrimination	0.00	0.01	0.03	0.37	0.71
Avg. Usefulness	-0.06	.02	-0.31	-3.35	0.00***
Avg. Interest	0.05	.02	0.29	3.08	0.00***

DV: Arguments, $R^2 = 0.26, Adj. R^2 = 0.20$

* $p < .05$, ** $p < .01$, *** $p < .001$

The model using demographic factors, learner characteristics, and multiple source use behaviors in predicting SOLO scores was significant, $F(12, 172) = 5.37, p < 0.001, R^2 = 0.27$. A model summary is presented in Table 32.

Table 32

Demographics, Learner Characteristics, and Multiple Source Use Behaviors Predicting SOLO Scores

Variable	B	Std. Error B	β	t	Sig.
<i>Step 1: Demographic Factors</i>					
Gender (1=Female)	0.17	0.17	0.07	1.02	0.31
Minority Status (1=Non-White)	-0.46	0.16	-0.20	-2.99	0.00*
Class Standing	-0.11	0.07	-0.11	-1.60	0.11
<i>Step 2: Learner Characteristics</i>					
Prior Knowledge	0.04	0.04	0.10	1.16	0.25
Attitudes	0.05	0.06	0.07	0.84	0.40
Source Eval Behaviors	-0.08	0.07	-0.09	-1.23	0.22
<i>Step 3: Multiple Source Use Behaviors</i>					
Total Source	0.03	0.07	0.04	0.49	0.63
Total Time	0.00	0.00	0.18	2.16	0.03*
Access Doc. Info	0.37	0.20	0.14	1.84	0.07
Trustworthiness Discrimination	0.01	0.00	0.12	1.51	0.13
Avg. Usefulness	-0.01	0.01	-0.13	-1.45	0.15
Avg. Interest	0.01	0.01	0.20	2.13	0.04*
DV: SOLO Scores, $R^2 = 0.27, Adj. R^2 = 0.22$					
* $p < .05$, ** $p < .01$, *** $p < .001$					

The model using demographic factors, learner characteristics, and multiple source use behaviors in predicting number of citations produced was significant, $F(12, 172) = 2.07$, $p < 0.05$, $R^2 = 0.13$. A model summary is presented in Table 33.

Table 33

Demographics, Learner Characteristics, and Multiple Source Use Behaviors Predicting Citations

Variable	B	Std. Error B	β	t	Sig.
<i>Step 1: Demographic Factors</i>					
Gender (1=Female)	0.39	0.32	0.09	1.24	.22
Minority Status (1=Non-White)	-0.71	0.29	-0.18	-2.46	0.02*
Class Standing	-0.09	.12	-0.05	-0.69	0.49
<i>Step 2: Learner Characteristics</i>					
Prior Knowledge	0.08	0.07	0.10	1.15	0.25
Attitudes	-0.004	0.12	-0.00	-0.04	0.97
Source Eval Behaviors	-0.10	0.13	-0.06	-0.76	0.45
<i>Step 3: Multiple Source Use Behaviors</i>					
Total Source	0.03	0.13	0.02	0.20	0.84
Total Time	0.00	0.00	-0.04	-0.45	0.65
Access Doc. Info	0.98	0.38	0.21	2.62	0.01**
Trustworthiness Discrimination	0.01	0.01	0.06	0.73	0.47
Avg. Usefulness	-0.01	0.01	-0.04	-0.44	0.66
Avg. Interest	0.01	0.01	0.08	0.77	0.44

DV: Citations, $R^2 = 0.13$, $Adj. R^2 = 0.07$

* $p < .05$, ** $p < .01$, *** $p < .001$

Appendix B: Demographics Questionnaire

1. First Name:
2. Last Name:
3. Email:
4. Course/Instructor (if applicable):

5. Age:
6. Gender:
 - a. Female
 - b. Male
 - c. Other
7. Race/Ethnicity:
8. GPA:
9. Major:
10. Class Standing:
 - a. Freshman
 - b. Sophomore
 - c. Junior
 - d. Senior
 - e. Other (please specify):

Appendix C: Prior Knowledge Identification Measure

Cronbach's $\alpha = 0.91$

Please tell me about each:

Mohamed Morsi:

Gender el-Sisi:

Muslim Brotherhood:

Arab Spring:

Tahrir Square:

Hosni Mubarak:

Tamarod:

Note: Terms and names were presented to participants in random order.

Appendix D: Stance Items

Who should the US support to hold power in Egypt?

- a) Mohamed Morsi and the Muslim Brotherhood
- b) General el-Sisi and the Military
- c) I don't know
- d) Other: _____

How strongly do you feel about events associated with the Arab Spring in Egypt?

Endorsed on a seven-point likert-scale ranging from 1 or "I have no opinion on the issue" to 7 or "I have very strong opinions on the issue"

Appendix E: Credibility Assessment Scale

How frequently do you do each of the following when looking for information for schoolwork:

1. Check to see if the information is current
2. Seek out other sources to validate the information
3. Consider whether the views represented are facts or opinions
4. Check to see who the author is
5. Check to see what other sites link to the Web site you are viewing
6. Check the qualification or credentials of the author
7. Visit the “about us” page on a Web site
8. Check to see if the information is complete/comprehensive
9. Consider the author’s goals/objectives for posting information

Students endorsed each item using a seven-point likert-scale ranking from 1 or “never” to 7 or “very often”

Appendix F: Academic Behaviors Questionnaire

Internet-Based Learning Activities (Bråten & Strømsø, 2006)

To what extent is each of the statements below typical of you:

1. When I use the Internet, I tend to spend time on materials that are not relevant
2. When I search for information on the Internet, I often feel like I'm drowning in information
3. I have difficulty identifying important information which I use the Internet
4. When I find information on the Internet that is relevant, it is often hard to assess its quality
5. When I use the Internet, I have difficulty finding information that I can use in my studies
6. When I use search engines like Google, I often have difficulty finding the information that I need
7. I have difficulty entering the right search terms when I am looking for information on the Internet

1 = not at all typical of me; 7 = very typical of me

Information Source Scale (Metzger et al., 2003)

How often do you use each of these information sources when completing schoolwork:

1. Journal articles
2. Public opinion surveys (e.g., Pew Research Center)
3. Newspapers
4. Textbooks
5. Wikipedia
6. Twitter
7. Blogs
8. Reports from research centers (e.g., Brookings)

1 = never; 5 = all the time

Information Source Credibility Scale (Metzger et al., 2003)

How credible do you consider each of these information sources to be:

1. Journal articles
2. Public opinion surveys (e.g., Pew Research Center)
3. Newspapers
4. Textbooks
5. Wikipedia
6. Twitter
7. Blogs
8. Reports from research centers (e.g., Brookings)

1 = not at all credible; 7 = very credible

Appendix G: Screenshot of Directions Page

In this study, you will be asked to provide a response to the following prompt:

Should the United States support General el Sisi and the military regime or Mohamed Morsi and the Muslim Brotherhood?

Directions:

Please answer as you would if assigned to write a brief essay in response to this prompt for an academic class.

In responding to the prompt you will be asked to **take a position** (i.e., in support of Mohamed Morsi or General el Sisi) as well as to **provide specific evidence to support your position**.

On the next page you will find a **source library** you can use to research information to help you in responding to the prompt.

1. You can use any of the sources in the library. You can take as much time as you need.
2. As you research, take notes in the open Word document or on the paper provided.
3. In composing your response, you will be able to use your notes **but will not be able to access any of the sources in the library**.

Appendix H: Screenshot of Response Window

Should the United States support General el Sisi and the military regime or Mohamed Morsi and the Muslim Brotherhood?

General El Sisi

Ex-President Morsi

Other

How did you come to this conclusion?

[Continue](#)

Appendix I: Texts in the Source Library

Analysis Essay

Title: El Sisi's Islamist Agenda for Egypt: The General's Radical Political Vision

Author: Prof. Robert Springborg, Council on Foreign Relations

Last week, el Sisi, the commander of Egypt's armed forces, which recently deposed the country's first freely elected president, went beyond addressing the graduates of Egypt's academies. El Sisi's true audience was the wider Egyptian public. He presented himself less as a general than as a populist strongman. He urged Egyptians to take to the streets to show their support for the provisional government that he installed after launching a coup to remove from power President Morsi, a longtime leader of the Islamist Muslim Brotherhood. El Sisi requested a "mandate" to confront the Muslim Brotherhood, whose supporters have launched protests to denounce the new military-backed regime.

El Sisi will not be content to serve as the leader of Egypt's military. There are indications that Sisi is less than enthusiastic about democracy and that he intends to hold on to political power. Based on a thesis he wrote at the U.S. Army War College it is possible that Sisi has something in mind: a hybrid regime that combines Islamism with militarism.

Sisi emphasizes the centrality of religion to the politics of the region, arguing, "for democracy to be successful in the Middle East," it must show "respect to the religious nature of the culture" and seek "support from religious leaders." He condemns governments that "disenfranchise large segments of the population who believe religion should not be excluded from government."

Indeed, Morsi tapped el Sisi to become defense minister because there was plenty of evidence that the general was sympathetic to Islamism. He is a devout Muslim who frequently inserts Koranic verses into conversations, and his wife wears the conservative dress favored by orthodox Muslims.

Those concerned about el Sisi's views on women's rights were alarmed by his defense of the military's use of "virginity tests" for female demonstrators. Human-rights activists argued that the "tests" amounted to sexual assaults.

Last summer, when Morsi tapped Sisi to replace the Minister of Defense, Morsi believed that he chose someone willing to subordinate himself to an elected government. Foreign observers interpreted Sisi's promotion as a signal that the military would be professionalized, with a reduction of its role in politics and the economy.

There is plenty of evidence that Sisi is not nearly as modest as he prefers Egyptians believe. He remained minister of defense in the new government and also took the post of first deputy prime minister. Sisi also released a video glorifying the general, taking care to illustrate the military's provision of goods and services to civilians.

He also tried to undercut support for the Brotherhood by appealing directly to their followers, these moves may be intended to inoculate him against the charge that the coup was anti-Islamist - a critical point, since Islamism enjoys broad support in Egypt.

If el Sisi continues to seek legitimacy for military rule by associating it with Islamism, it could be a disaster for Egypt, setting back the democratic cause. It would also reinforce the military's octopus-like hold on the economy, which is a major obstacle to Egypt's economic development.

Blog Post

Title: A Message to the Muslim Brotherhood: Enough Lies, Enough Blood

Author: The Idealist 11

I highly doubt that the police or the army is responsible for the deaths of the tens of Muslim Brotherhood members early this morning. A security official has said that only tear gas, not live ammunition, was used to disperse the Morsi supporters. On the other hand, the Brotherhood spokesman told Reuters news agency: "They are not shooting to wound, they are shooting to kill."

But ask yourselves. What would the security forces gain from killing members of the pro-Morsi protestors? Think rationally. Killing the protestors would actually work against them. The armed forces are currently widely popular among a large sector of the Egyptian population, which is against Morsi and Muslim Brotherhood rule. If the security forces were responsible for killing peaceful Morsi supporters, they would be highly discredited and their popularity would be seriously shaken.

So now we have to ask ourselves: who would benefit from this? The answer is simple. The Muslim Brotherhood. It would work in their favour to demonize the army.

There is NO evidence that the armed forces were the ones shooting at the protestors. General Sisi promised to protect peaceful protests irrespective of their affiliation and only use violence in self-defense against violent protestors. His accountability would be shattered if the armed forces were behind the deaths of peaceful protestors. And that is exactly what the Muslim Brotherhood want.

Not only is it in their interest, but their own leaders have been calling on them to commit jihad against the "military coup." They would sacrifice themselves and kill their own people to demonize the army. Ex-Muslim Brotherhood members have confirmed this. I'm not saying that they shot their own people in this specific incident. But they are acting provocatively and as the interior ministry spokesman said "They are purposely causing a crisis." In his statement he said the Morsi supporters: "halted traffic, set tires on fire and clashed with residents of the nearby [working class] Mansheyet Nasr district using live fire and birdshots, and this killed 21 people."

So think before blindly believing everything you read. The Muslim Brotherhood members and other Islamist groups are religious fanatics that will resort to any measures to reach a desired end. I am not generalizing this assertion to all of the pro-Morsi demonstrators. However, I am especially critical of Islamist groups that use radical rhetoric and the name of God to justify violence and jihad.

News Paper Article

Title: Egypt Arrests 11 Islamists for Facebook Activity

Author: Maggie Michael, Associated Press

Egyptian authorities, led by General el Sisi, arrested 11 Muslim Brotherhood members accused of running Facebook pages that incited violence against police.

After crippling the Muslim Brotherhood in a wave of arrests and killings of protesters, security agencies are going after younger supporters of ousted Islamist President Morsi, looking for links to a growing violent backlash against the military regime. Younger members have been using the Internet and social media to keep Islamist protests alive.

Bombings and drive-by shootings targeting police officers have accelerated in retaliation for the killings and jailing of Brotherhood members and Islamists. The government accuses Morsi's Brotherhood of orchestrating the violence and has branded it a terrorist group — an accusation the Brotherhood denies.

Police pursuit of protestors online raises concerns that attacks could be used as a pretext for imposing heavier restrictions on the freedom of the Internet, which has served as a major outlet for protests after the military-backed interim government of General el-Sisi silenced other forms of dissent, arresting secular activists.

Social media has become the Brotherhood's main form of communication, spreading calls for non-stop protests demanding Morsi's reinstatement and posting pictures of slain protesters while calling for revenge.

After police crashed pro-Morsi protest camps, Brotherhood websites were filled with pictures and personal details about police officers involved in the assault. The day witnessed one of Egypt's worst bloodbaths, with hundreds killed.

The arrests today were linked to Brotherhood websites. The Interior Ministry said those detained were accused of using websites to "incite violence, target citizens, make bombs and carry threatening messages."

Among those detained were six people running a Facebook page, "Damanhour Ghosts," with around 500 followers, criticizing the military government and calling for the freeing of detainees.

Others were arrested in connection to a page, "Anti-Coup Hooligans Brigade," which includes photos of policemen accused of killing protesters, warning, "vengeance is coming" and pictures of youth throwing firebombs.

During Morsi's one-year presidency, there were several arrests related to postings considered blasphemous or offensive to Islam. Coptic Christian Saber was sentenced to prison after neighbors complained that he posted an anti-Islamic film on his Facebook page. In 2012, blogger and activist Abdel-Fattah was detained after a post on his Twitter account.

Currently, under General Sisi, authorities have expanded their anti-Brotherhood crackdown to go after the Qatar-based Al-Jazeera TV news network, which they have long accused of being biased toward the Brotherhood. The network denies any bias.

Twenty Al-Jazeera journalists were put on trial for aiding a terrorist group and endangering national security.

U.S. State Department spokeswoman Psaki said Washington was "deeply concerned" about the lack of freedoms in Egypt and the country's "egregious disregard for the protection of basic rights and freedoms."

Egypt's Foreign Ministry spokesman, Abdelattie, rejected the U.S.'s criticism, insisting that the judicial system ensures fair trials with no government interference.

No date has been set for the trial and the full list of charges and defendants has not been released.

Public Opinion Poll

Title: Egyptian Attitudes

Author: Dr. James Zogby and Associates

This poll represents the third time since May 2013 that we have conducted a nationwide survey of Egyptian public opinion. Our polling has allowed us to track the Egyptians' changing attitudes toward developments as they unfold: how they view their government and institutions; areas of agreement and disagreement; and their hopes for the future.

In our September survey, we find that public opinion in Egypt has become more conflicted and more polarized. Overall, 60% of Egyptians remain hopeful about the country's future and 83% believe that the situation will improve in the next few years, but the continuing violence has taken a toll. A plurality (46%) of all Egyptians believe that the situation in their country has become worse, not better, since the Morsi government was deposed. Eighty percent (80%) of Morsi's Freedom and Justice Party (FJP) supporters express this view. But only about one-half of the rest of the country feels that Egypt is better off, with nearly one in five saying that the situation is the same as it was before the military intervened.

The military remains the institution in which Egyptians have the greatest confidence, but their positive rating has declined to 70%, owing to a sharp drop in support from those who identify with the Muslim Brotherhood's FJP and a slight decline in support among liberals.

The country is split down the middle in its view of the military's July 3rd deposing of the Morsi government. The FJP, of course, is unanimous in finding the military's action incorrect, while almost two-thirds of the rest of Egyptians support the deposing of Morsi.

Between July and September, confidence in the interim government of el Sisi has increased, with between 43% and 51% now saying that they believe that this government can follow the "roadmap" and restore order to the country – with almost two-thirds of non-FJP supporters now expressing this view.

During the past month, the Muslim Brotherhood's party has consolidated its strength, while at the same time alienating itself from many other Egyptians. Support for the FJP has leveled off at 34%, up from May's 28%. And 79% of all Egyptians still want national reconciliation as the desired goal for Egypt. But now one-half of those who do not support FJP identify the Muslim Brotherhood as the main obstacle to reconciliation and more than 60% of non-FJP supporters want the Brotherhood to be banned from Egyptian politics.

With the FJP continuing to have the support of about one-third of the country, some effort to achieve national reconciliation will be important. At this point, it appears that the choices made by both the military and the Muslim Brotherhood will be decisive in shaping Egypt's near-term future.

Twitter

Account: @Ikhwanweb, English language Twitter account of the Muslim Brotherhood

Ibrahim Yehia Azab, #AntiCoup pharmacy student, will be executed Tuesday on trumped up charges without due process

We hold ruling military regime fully responsible for the loss of innocent Egyptian lives; for failure to protect security and rule of law.

Military regime always rush,for pol reasons,to implicate MB in these crimes w/o a shred of evidence & before any investigation is launched

Targeting of soldiers and/or civilians is heinous crime that requires thorough & transparent investigation to bring perpetrators to justice

Muslim Brotherhood Statement Marking the #IWD2014: Egyptian Women Suffer Inhuman Coup Crimes

Students protests erupt across university campuses calling for freedom, democracy & end to oppressive military rule

Brutality, torture, rape: Egypt's crisis will continue until military rule is dismantled

Human Rights Activist Khaled Hamza Detained at Unknown Location by Coup Forces

Muslim Brotherhood strongly condemn killing of a police sergeant in Mansoura Fri, reaffirm unequivocal rejection of all forms of violence

@EgyAntiCoup: BREAKING: Abdelrahman Mosadaq,16, killed by a live bullet to the neck after police opened fire at a protest in Alexandria"

How did 37 prisoners come to die at Cairo prison Abu Zaabal?

Statement: Muslim Brotherhood Believes in Popular Will, Democracy, Despite Junta Attempts to Eradicate Group

Human Rights Monitor Condemns Violence Against Women Opposed to the Coup in Egypt

Two Egyptians killed, tens injured as a result of police attacks on peaceful #AntiCoup protests throughout the country Friday

Massive #AntiCoup protest in Matariya.Worth mentioning that scores were killed in Matariya on #Jan25.

Thank u to @mosaaberizing & all photojournalists for documenting the massacre so we don't forget this day

#AntiCoup protests continue across Egypt, calling for freedom, and an end to authoritarianism and military rule

Illegal and inhumane conditions of anti- coup detainees in El- Akrab jail. #Egypt #AntiCoup #Human_Rights

PR campaign to undermine #Anticoup movement,force opposition to accept the Generals' roadmap rejected by marjority of Egyptians

@EduardCousin Yes. it's untrue. Our position is clear; Morsi is the legitimate prez of Egypt, military must not have any role in politics

@TheCairoPost You keep posting false news about us, but this one is outrageous. Please verify before you publish

Wikipedia Entry

Title: 2013 Egyptian Coup d'état

Selected Editors: Omar Othman 95, GreysHark09, Alhanuty

On July 3, 2013, the Egyptian army chief General el-Sisi removed the country's then-incumbent president Mohamed Morsi from power and suspended the Egyptian constitution. The move came after four days of mass revolt and it also came 48 hours following a warning from the army to respond to the demands of the millions of protesters or it would intervene to restore order and impose its own roadmap. Al-Sisi declared Chief Justice of the Supreme Constitutional Court of Egypt Mansouras the interim president of Egypt. Morsi was put under house arrest and several Muslim Brotherhood leaders were arrested. The announcement was followed by demonstrations and clashes between supporters and opponents of the move throughout Egypt. The announcement was followed by a statement in support of the military's action by the Grand Sheikh of Al Azhar Ahmed el-Tayeb, the Coptic Orthodox Pope Tawadros II as well as opposition leader Mohamed El Baradei.

There were mixed international reactions to the events. Most Arab leaders were generally supportive or neutral, with the exception of Qatar and Tunisia who strongly condemned the military's actions. Other states either condemned or expressed concern over the removal of Morsi; there was also a perceived measured response from the United States. Because Egypt had an interruption of constitutional rule, Egypt was suspended from the African Union. There has also been debate in the media regarding the labeling of these events. It has been variously described as a coup d'état or as a revolution, by proponents. Ensuing protests in favour of Morsi were violently suppressed with the dispersal of pro-Morsi sit-ins on August 14, amid ongoing unrest.

The removal of Morsi from office by the military was a result of a coup d'état following protests, that were instigated by frustration with Morsi's year-long rule. Issues that might have lead to the removal of Morsi by the army include:

- Morsi issuing a constitutional declaration that made all of his decisions binding and unable to be appealed until the constitution was approved and a new People's Assembly elected. This was later abrogated due to multiple protests and the public's anger.
- Muslim Brotherhood using their majority to pass a hotly disputed constitution. Though the constitution was approved by 64% of Egyptians, only about 30% of Egyptians with voting rights participated in the referendum process. Secular and Liberal Party members and church representatives withdrew from the constitutional committee.
- Power, gas and economic crises.
- Diplomatic problems including construction of an Ethiopian dam along the Nile River, affecting Egypt's share of water.
- Security of the state worsened; two of the most prominent stories related to security under Morsi were: the murder of 16 border guards in Sinai and the abduction of 7 Egyptian security personnel who were later released.
- Egyptian Army's economic interests. These include foreign partnerships related to maritime and air transport, oil and gas and industrial-scale environmental projects such as wastewater treatment and energy generation. Modern economic development has always threatened the military's economic position.

Appendix J: Screenshot of Source Library

Please use the source library below to research information to respond to the following prompt:

Should the United States support General el Sisi and the military regime currently in power in Egypt or Mohamed Morsi and Muslim Brotherhood protestors?

Please take notes while you research. You will not be able to refer to the sources while composing your response.

Blog Post

Newspaper Article

Twitter

Wikipedia Entry

Public Opinion Survey

Analysis Essay

[Click here to continue](#)

Appendix K: Screenshots of Accessible Document Information

Newspaper Article ✕

[Click here to learn more about this source](#)

Egyptian authorities, led by General el Sisi, arrested 11 Muslim Brotherhood members accused of running Facebook pages that incited violence against police.

After crippling the Muslim Brotherhood in a wave of arrests and killings of protesters, security agencies are going after younger supporters of ousted Islamist President Morsi, looking for links to a growing violent backlash against the military regime. Younger members have been using the Internet and social media to keep Islamist protests alive.

Bombings and drive-by shootings targeting police officers have accelerated in retaliation for the killings and jailing of Brotherhood members and Islamists. The government accuses Morsi's Brotherhood of orchestrating the violence and has branded it a terrorist group — an accusation the Brotherhood denies.

Police pursuit of protestors online raises concerns that attacks could be used as a pretext for imposing heavier restrictions on the freedom of the Internet, which has served as a major outlet for protests after the military-backed interim government of General el-Sisi silenced other forms of dissent, arresting secular activists.

Newspaper Article ✕

Title: Egypt Arrests 11 Islamists for Facebook Activity
Author: Maggie Michael
Publisher: Associated Press
Date of Publication: January 30, 2014
Location: Cairo, Egypt
URL: <http://bigstory.associatedpress.org/>

Author Information: Maggie Michael, Associated Press Senior Reporter, Cairo Bureau

Egyptian authorities, led by General el Sisi, arrested 11 Muslim Brotherhood members accused of running Facebook pages that incited violence against police.

After crippling the Muslim Brotherhood in a wave of arrests and killings of protesters, security agencies are going after younger supporters of ousted Islamist President Morsi, looking for links to a growing violent backlash against the military regime. Younger members have been using the Internet and social media to keep Islamist protests alive.

Bombings and drive-by shootings targeting police officers have accelerated in retaliation for the killings and jailing of Brotherhood members and Islamists. The government accuses Morsi's Brotherhood of orchestrating the violence and has branded it a terrorist group — an accusation the Brotherhood denies.

Twitter



Author: @Ikhwanweb
Publisher: Twitter
Dates of Publication: February 14, 2014 – February 17, 2014
Location: London, UK
URL: <https://twitter.com/Ikhwanweb>

Author Information: @Ikhwanweb, Official English language Twitter account of the Muslim Brotherhood

Ibrahim Yehia Azab, #AntiCoup pharmacy student, will be executed Tuesday on trumped up charges without due process

We hold ruling military regime fully responsible for the loss of innocent Egyptian lives; for failure to protect security and rule of law.

Military regime always rush,for pol reasons,to implicate MB in these crimes w/o a shred of evidence & before any investigation is launched

Appendix L: Screenshot of Source Rating Window

Source Evaluation ×

Please rate this source:

How **useful** was this source?
Not at all | | Very

How **trustworthy** was this source?
Not at all | | Very

How **accurate was the information** in this source?
Not at all | | Very

How **interesting** was this source?
Not at all | | Very

[Back to Library](#)

Appendix M: Sample Response Coding

Response	No. Arguments	SOLO Score	Explanation	No. Cites
I decided in US would be better to go other because neither of the stories from the sources add up. Because none of the sources were a direct interview to get first hand evidence I don't have enough information to make a good choice.	0	0	No reasons provided	0
(1) After reading several statements from these various forms of media, I have come to the conclusion that General el Sisi is somewhat of a tyrant who needs to be stopped.	0	0.5	Attempting to provide a single reason; however missing justification	0
(1) They should support General Sisi because he is arrested the Muslim brotherhood members that have committed crimes throughout the country.	1	1	Single reason provided	0
The United States should support Ex-president Morsi. From the research I gathered, he is more in line with the desires/wishes of the people of Egypt. (1) Egyptians are being harassed, beaten up, killed because they are not able to express their political viewpoints under General el Sisi. And this notion of freedom of the people being undermined by the government does not coincide with US policies.	1	1.5	Attempting to provide multiple arguments; however not all are articulated fully	0
(1) According to the essay , General el-Sisi is not in favor of democracy and plans to hold power indefinitely. (2) Some sources claim that el-Sisi has committed terrorist acts and is overthrowing the president. The USA should not support a military regime that overthrows their own government.	2	2	Multiple reasons provided but listed, rather than integrated	1
(1) The newspaper article says that Morsi's Brotherhood is responsible for the deaths of numerous police officers. It is considered to be revenge for killing Brotherhood members. By offending Islam, one can expect retaliation from Morsi's Brotherhood, (2) and the newspaper article specifically mentions two cases where people were jailed for displaying their views on social media, whether it be Facebook or Twitter. (3) The blog post states that Morsi's Brotherhood is dangerous and that General Sisi is trying to protect the Egyptian citizens. Violence is not the answer, according to the blog author , and I whole-heartedly agree with his stance. (4) According to the public opinion survey , most Egyptians want to oust Morsi from his position of power. (5) With all the corruption and violence occurring in Egypt, the	7	2.5	Multiple reasons provided; attempting to integrate response as pro-Morsi, however, is more of a listing of evidence.	6

United States should do everything in its power to support general Sisi and free the citizens of Egypt from all these problems.

(6) The **analysis essay** shows that while general Sisi is a very religious person, he believes that the government should not mistreat its citizens by any means. People in the United States should realize how important religion, particularly Islam, is throughout the Middle East.

After searching diligently for answers in which to choose a side in this matter, the sources that I read allowed me to reach the conclusion that (7) Morsi is the bad guy in all of this. Too many people are dying, and corruption is running rampant throughout all of Egypt. The citizens are in dire need for a change, and the United States can help oust Morsi from his throne, so to speak.

With the current situation going on in the Middle East, the U.S. should be supporting the struggle to reinstate ex-president Morsi of Egypt. (1) According to the **report "El Sisi's Islamist Agenda for Egypt: The General Radical Political Vision" by Robert Springborg published by the Council on Foreign Relations**, Morsi, the countries first elected president, was ousted by General El Sisi who now is identified as not only the Minister of Defense but the First Deputy Prime Minister. (2) Since gaining power, Sisi has been in conflict with the Muslim Brotherhood who call for the reinstatement of the ex-president. (3) The **report** points out that being both head of state and military could lead to Sisi enacting what he wrote his thesis on in a U.S. military school: "a hybrid regime that combines Islamism with militarism." (4) The U.S. is already concerned with Sisi's handling of the Brotherhood protesters. According to the **Associated Press article written by Maggie Michael, "Egypt Arrests 11 Islamist for Facebook Activity"**, the U.S is "deeply concerned" about the lack of freedom in Egypt at the moment. (5) The U.S. needs to support Morsi, as there is already evidence of violence: there are images online of slain protesters with cries for revenge. Overall, with the amount of violence and freedom loss in the state of Egypt right now, it is imperative that the U.S. support the reinstatement of Morsi, Egypt's first elected president.

5

3

Integrated response supporting ex-president Morsi

3

(1) General El Sisi, although not democratically elected, (2) has the support of the people according to the **public opinion poll**. (3) According to the **blog post**, General El Sisi has promised not to use violence against peaceful protesters and as a country that supports the right of others to protest their leaders, the US should support this leader. However, after reading these six sources, I did not clearly know enough about Ex-President Morsi to completely determine the character of his rule, (4) although I do know from the **public**

7

3.5

Presents multiple arguments and attempts to compare and evaluate Morsi's and el Sisi's position, however points of comparison are not

6

opinion survey that the military was more strongly supported. (5) Additionally, as the **essay** purported, General El Sisi appointed someone else as president, so he is not exactly engaged in a traditional coup for power. (6) However, as evidenced by the **public opinion survey**, neither leader seems to be doing good for the country. A plurality think that the country has gotten worse and many believe nothing has changed at all. (7) However, due to the atrocious lack of freedoms in Egypt listed in the **Wikipedia article**, I am more inclined to think that the country is better under the rule of General El Sisi than under Ex-President Morsi.

(1) Though the Muslim Brotherhood seems to be a bit disorganized in terms of creating protests that do more than startle their enemies, I believe their cause is better for Egyptian citizens. (2) When Morsi was in power he created a constitution. Even though not all people agreed with what it said, it laid the ground works for creating a democracy and more fundamentals for a stable leadership. Their intent is to create a democracy, freedom, (3) and to end an oppressive military rule that frequently throws human rights to the side.

(4) General El Sisi does not have a strong intention for democracy and (5) believes more in holding political power to oppose all that go against him and his regime than for acting for the good of the people. (6) Binding himself and the country to religious standards and (7) using military to strictly enforce his religious beliefs does not make for a good leader. (8) Disallowing citizens to speak freely, even on the internet about their opinions when they oppose his views (9) by using violence and imprisonment as punishment and threat shows more intent on driving his own religious agenda forward instead of what the public wants or needs. (10) His abuse of power has allowed for too many human rights to be violated including sexually assaulting women to make a religious point about the sacredness of virginity is despicable, and should not be reinforced by supporting such a self-righteous leader.

(11) The United States should be supporting a leader that has the intention to create a country that cares about the opinions, safety and growth of its people through a process that is as peaceful as possible. The Muslim Brotherhood, though it is violent in its protests is the lesser of two evils, and has presented its interest in democracy and the production of constitutional documents in the past proving their hope for a brighter future.

entirely clear

11	4	Presents multiple reasons and an integrated response, evaluating and juxtaposing Morsi's and el Sisi's use of violence and intentions to create a democracy	0
----	---	---	---

Appendix N: Screenshot of Source Ranking and Justification Page

Please rank each source you used from **most useful** to **least useful**.

	Rank the source you consider to be most useful as 1	Please justify your ranking of each source.
Public Opinion Survey	1 <input type="radio"/> 2 <input type="radio"/>	<input type="text" value="Please enter your reasoning"/>
Twitter	1 <input type="radio"/> 2 <input type="radio"/>	<input type="text" value="Please enter your reasoning"/>

Please rank each source you used from **most trustworthy** to **least trustworthy**.

	Rank the source you consider to be most trustworthy as 1	Please justify your ranking of each source.
Public Opinion Survey	1 <input type="radio"/> 2 <input type="radio"/>	<input type="text" value="Please enter your reasoning"/>
Twitter	1 <input type="radio"/> 2 <input type="radio"/>	<input type="text" value="Please enter your reasoning"/>

Please rank the **information** in each source you used from **most accurate** to **least accurate**.

	Rank the source you consider to have the most accurate information as 1	Please justify your ranking of each source.
Public Opinion Survey	1 <input type="radio"/> 2 <input type="radio"/>	<input type="text" value="Please enter your reasoning"/>
Twitter	1 <input type="radio"/> 2 <input type="radio"/>	<input type="text" value="Please enter your reasoning"/>

Appendix O: Post-Task Engagement Scale

1. How much did you like completing the task?
2. How engaged were you in this task?
3. How interesting was this task?
4. How difficult was this task?
5. How much effort did you put into completing this task?

Students endorsed each item on a 7-point likert scale ranging from “not at all” to “very”. For instance, the engagement item ranged from, “not at all engaged” to “very engaged”. The interest item ranged from “not at all interesting” to “very interesting”.

Table 34

Item-Specific Reliability Statistics

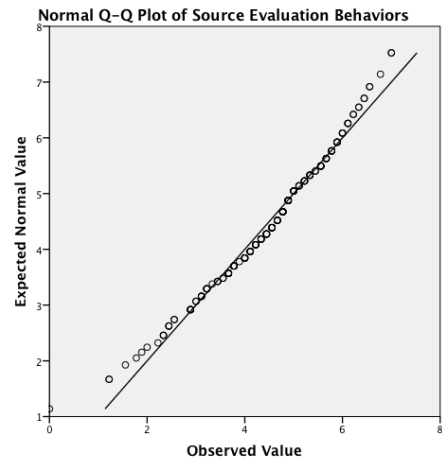
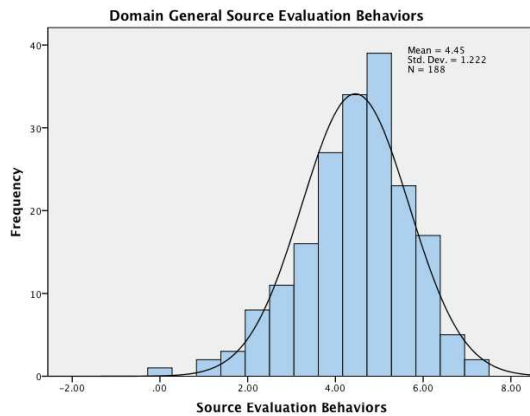
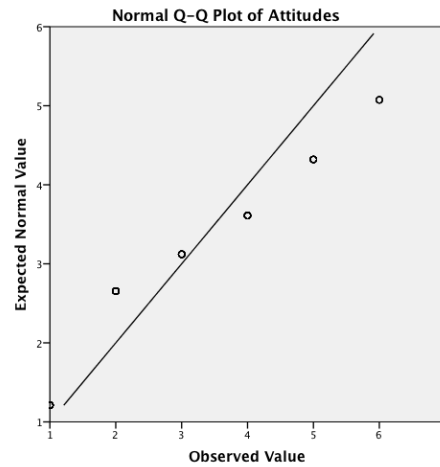
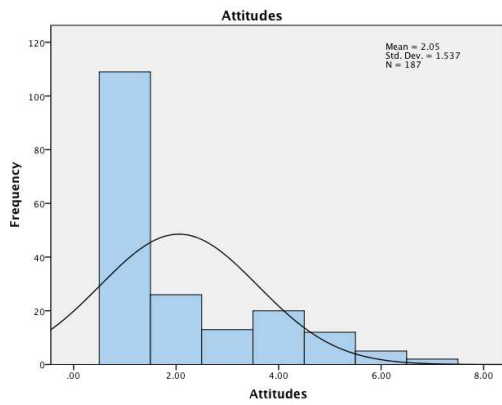
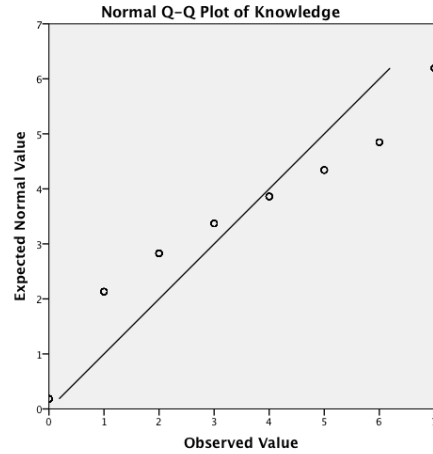
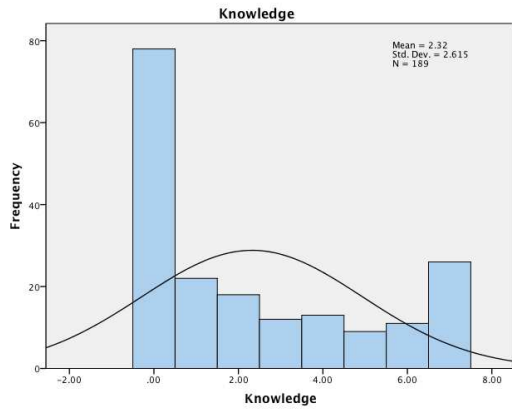
Item	Item-Total Correlation	Alpha if Item Deleted
Like	0.60	0.60
Engaged	0.64	0.60
Interesting	0.71	0.54
Difficult*	0.09	0.81
Effort	0.40	0.69

Cronbach's $\alpha = 0.71$ for 5 items

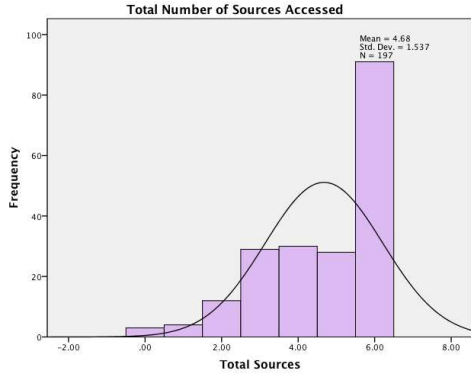
Note: The difficulty item was reverse coded.

Appendix P: Ascertaining Normality of Learner Characteristics and Multiple Source Use Behavioral Variables

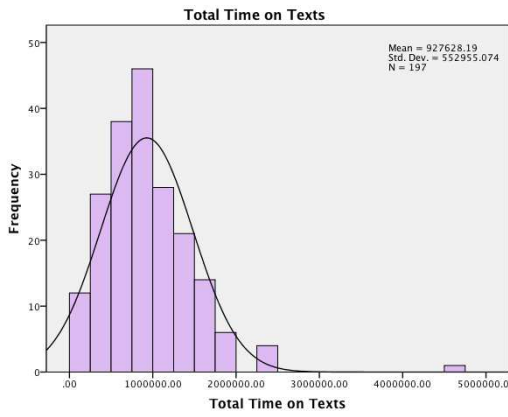
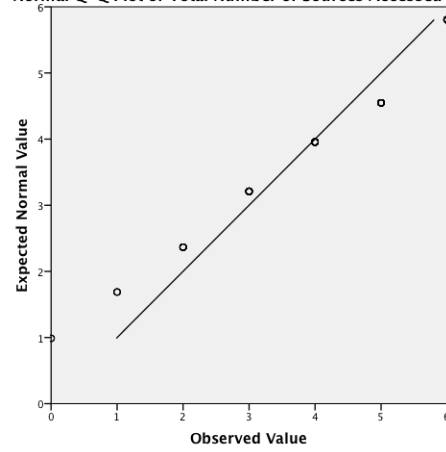
Learner Characteristics:



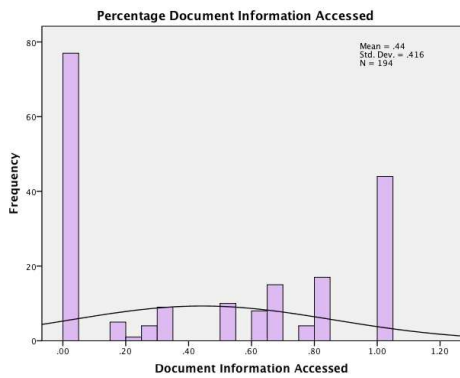
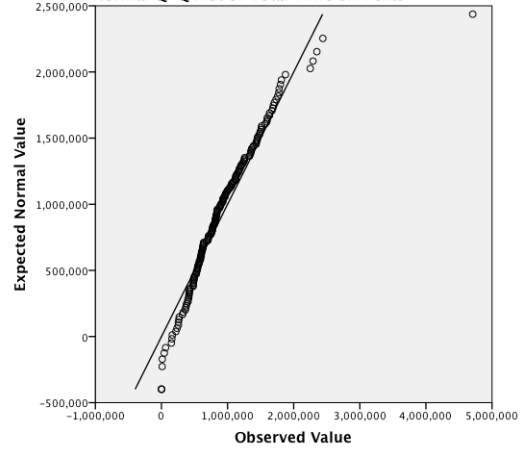
Multiple Source Use Behaviors:



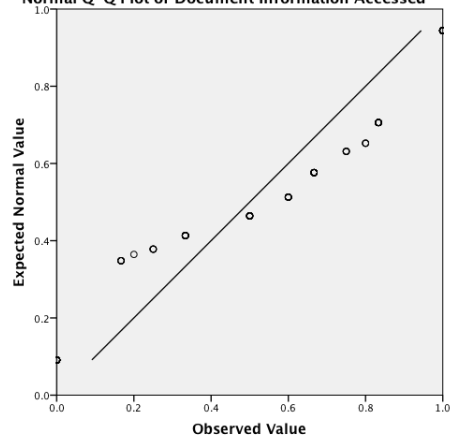
Normal Q-Q Plot of Total Number of Sources Accessed



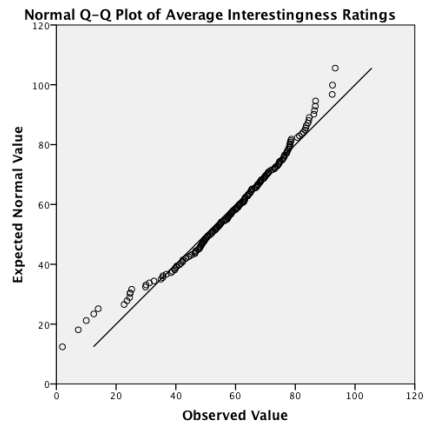
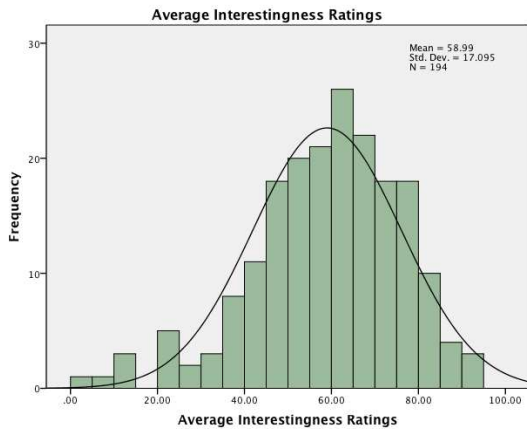
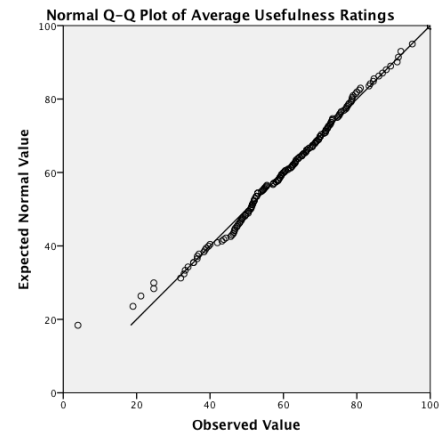
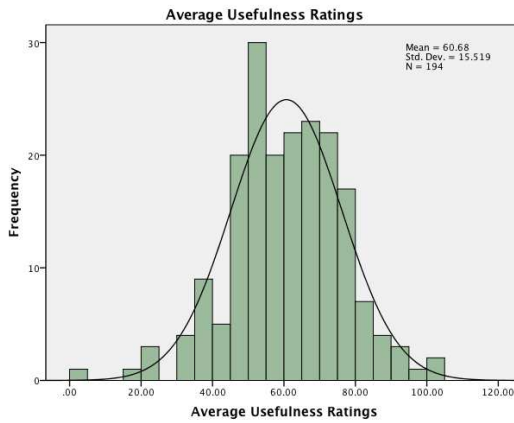
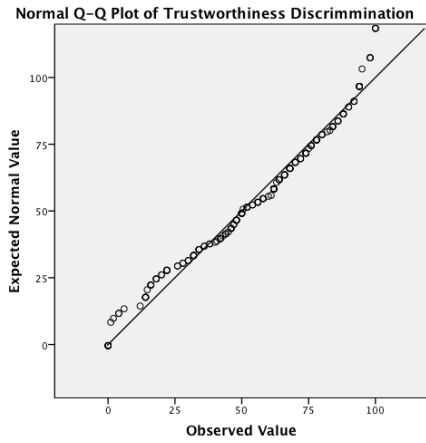
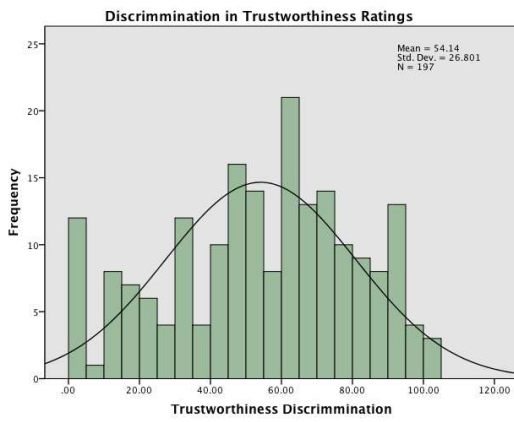
Normal Q-Q Plot of Total Time on Texts



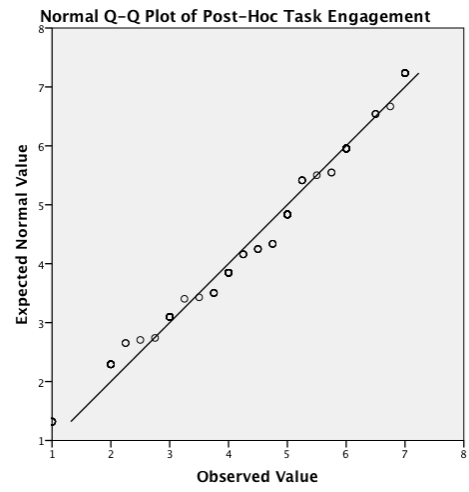
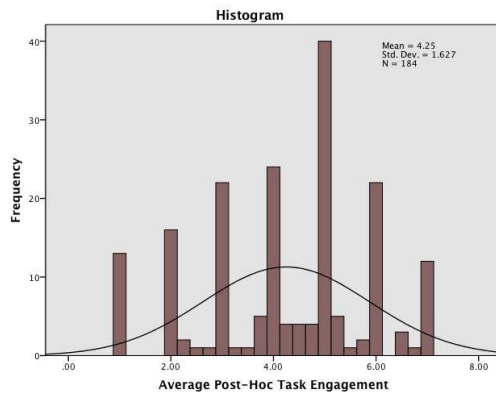
Normal Q-Q Plot of Document Information Accessed



Source Ratings:

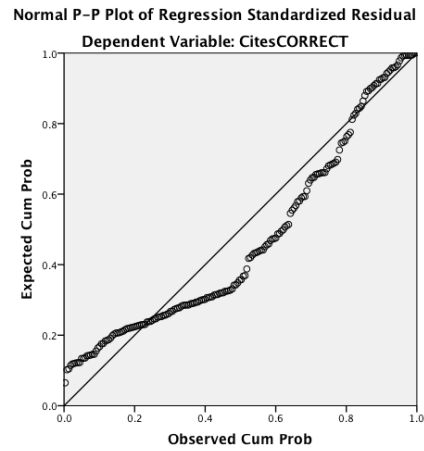
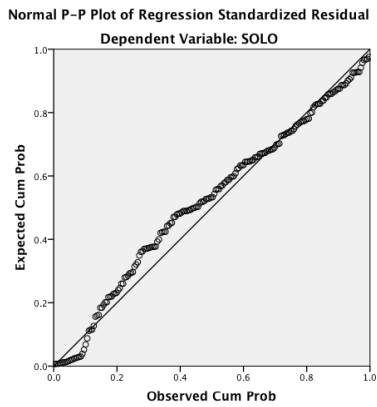
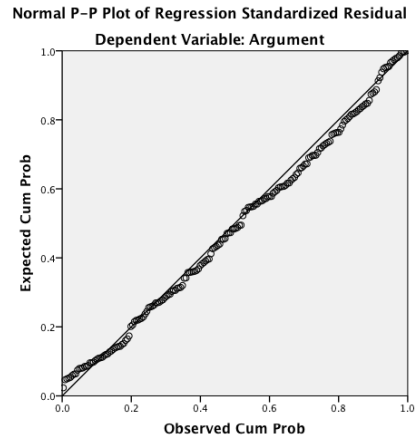
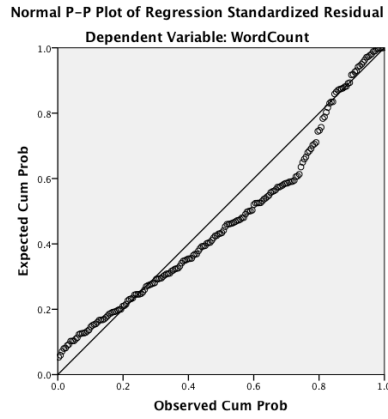


Task Engagement:

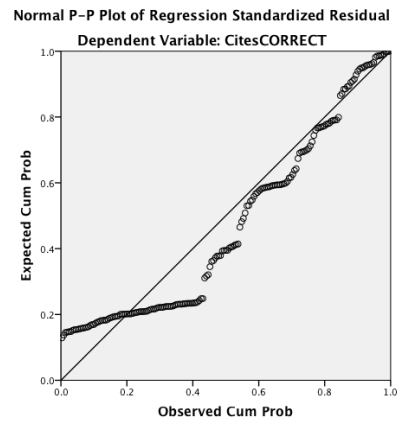
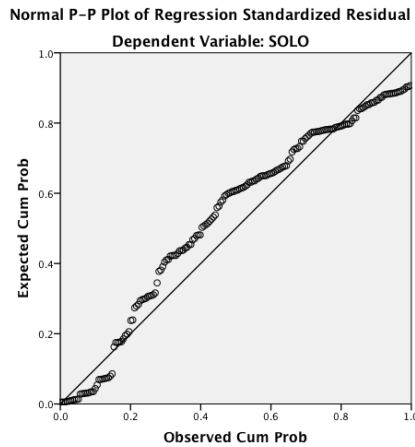
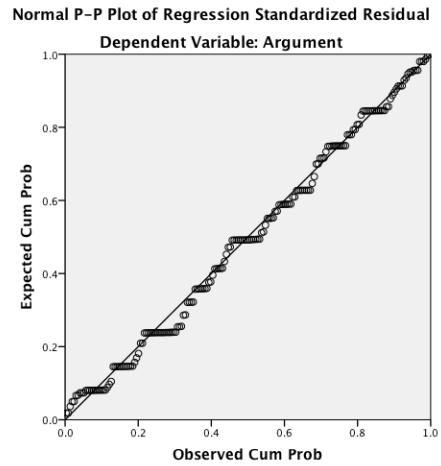
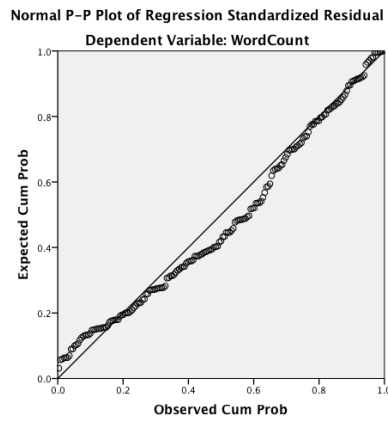


Appendix Q: Residual P-P Plots for Predictive Models

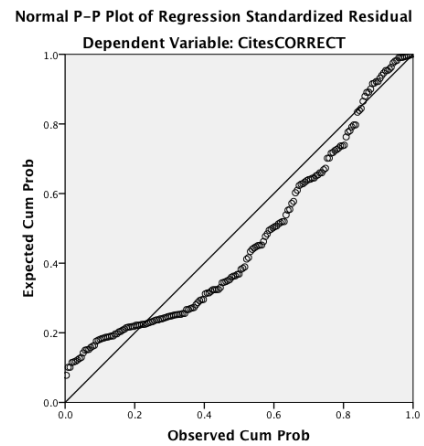
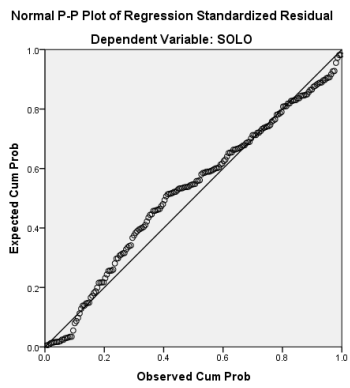
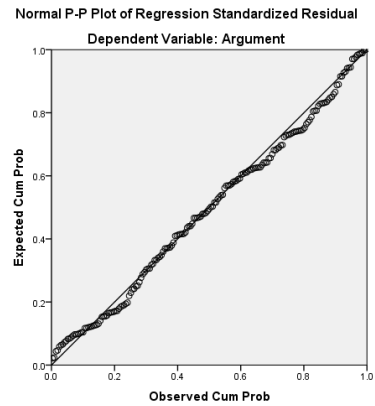
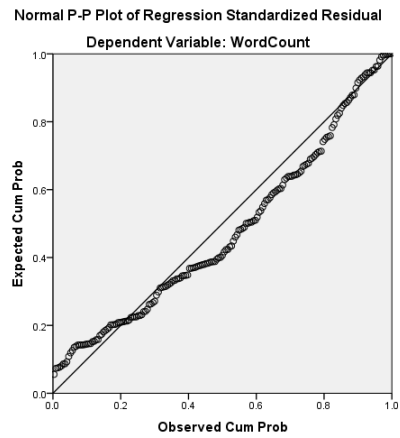
P-p plots using MSU behaviors to predict each outcome variable



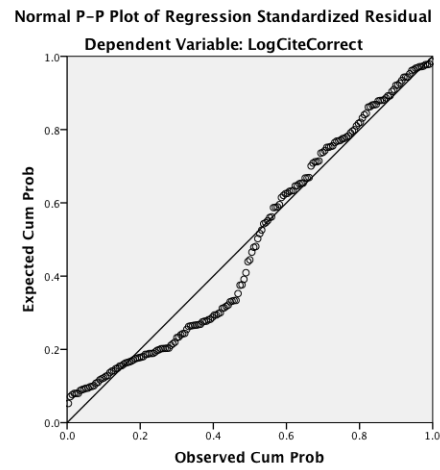
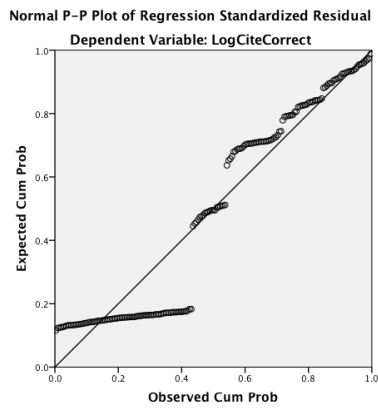
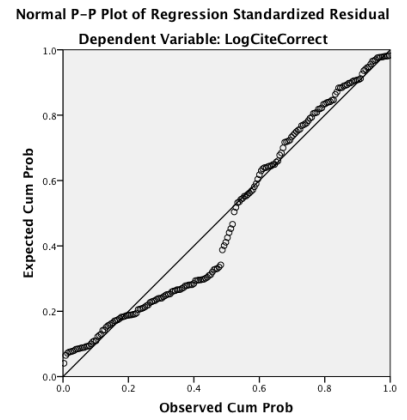
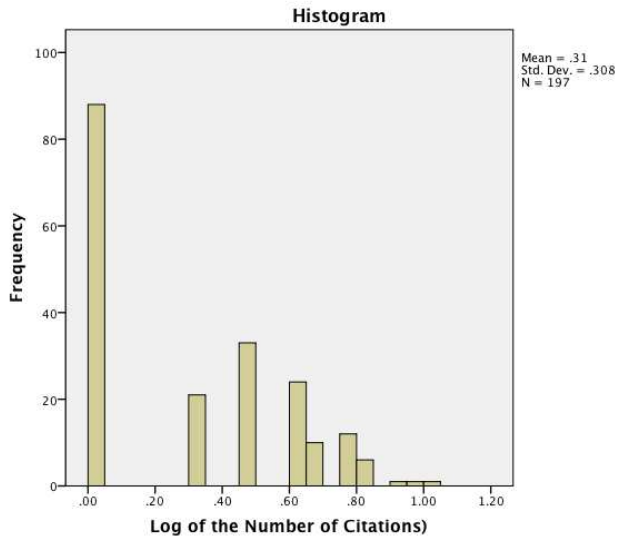
P-p plots using learner characteristics to predict each outcome variable.



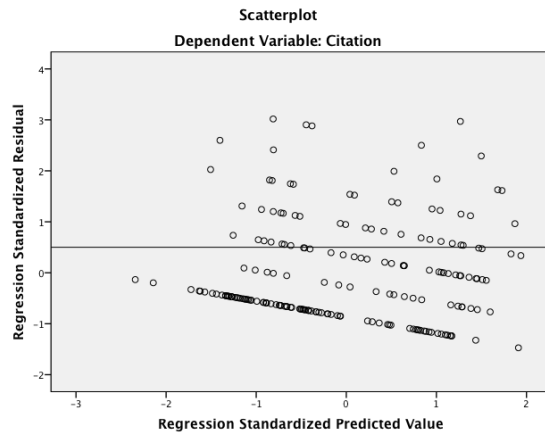
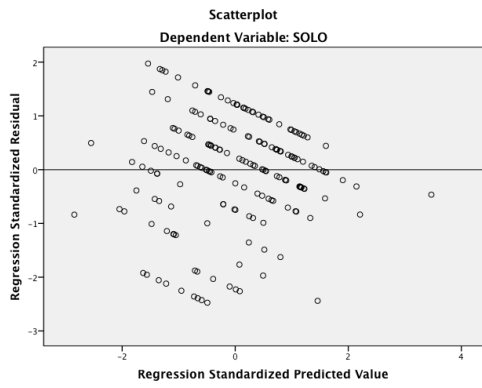
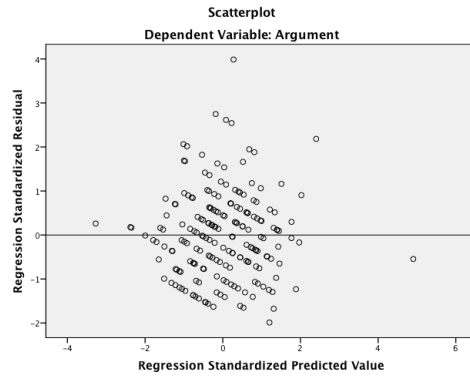
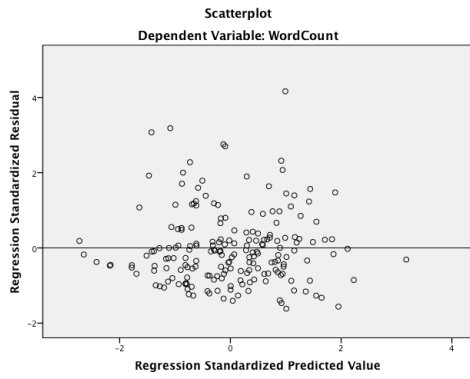
P-p plots using learner characteristics and MSU behaviors to predict each outcome variable



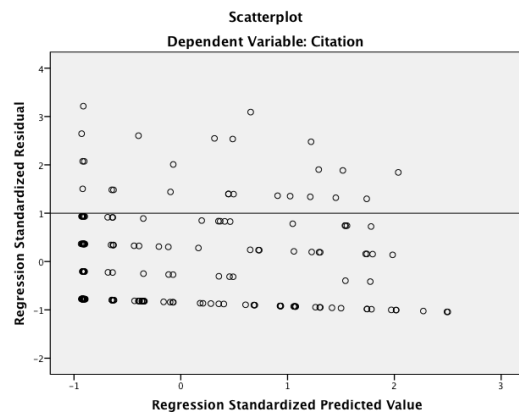
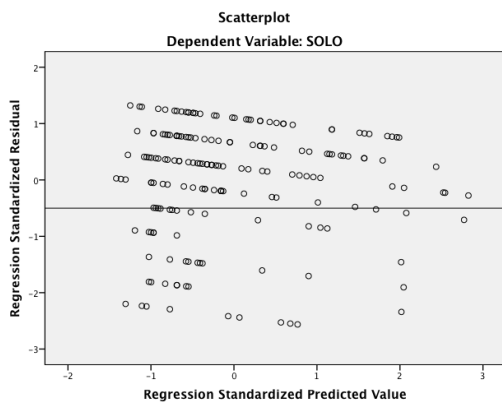
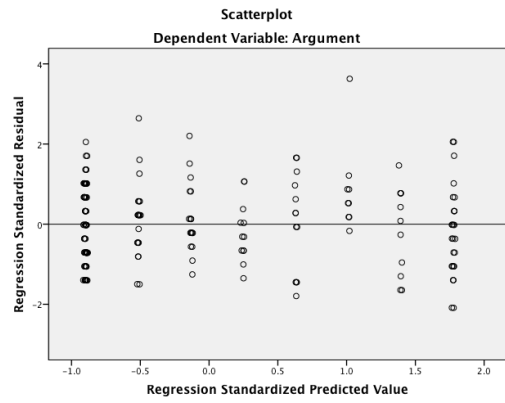
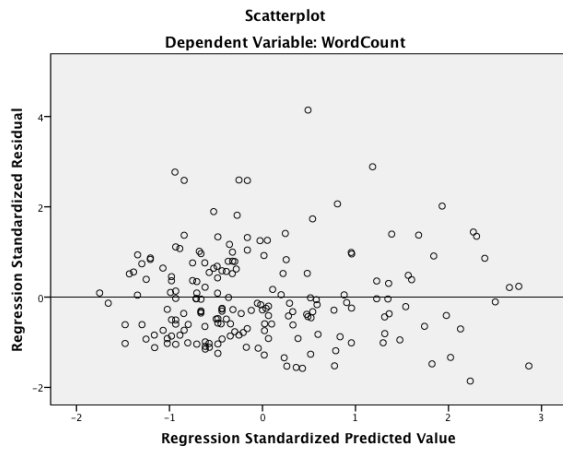
Histogram and residual p-p plot of transformed citation variable



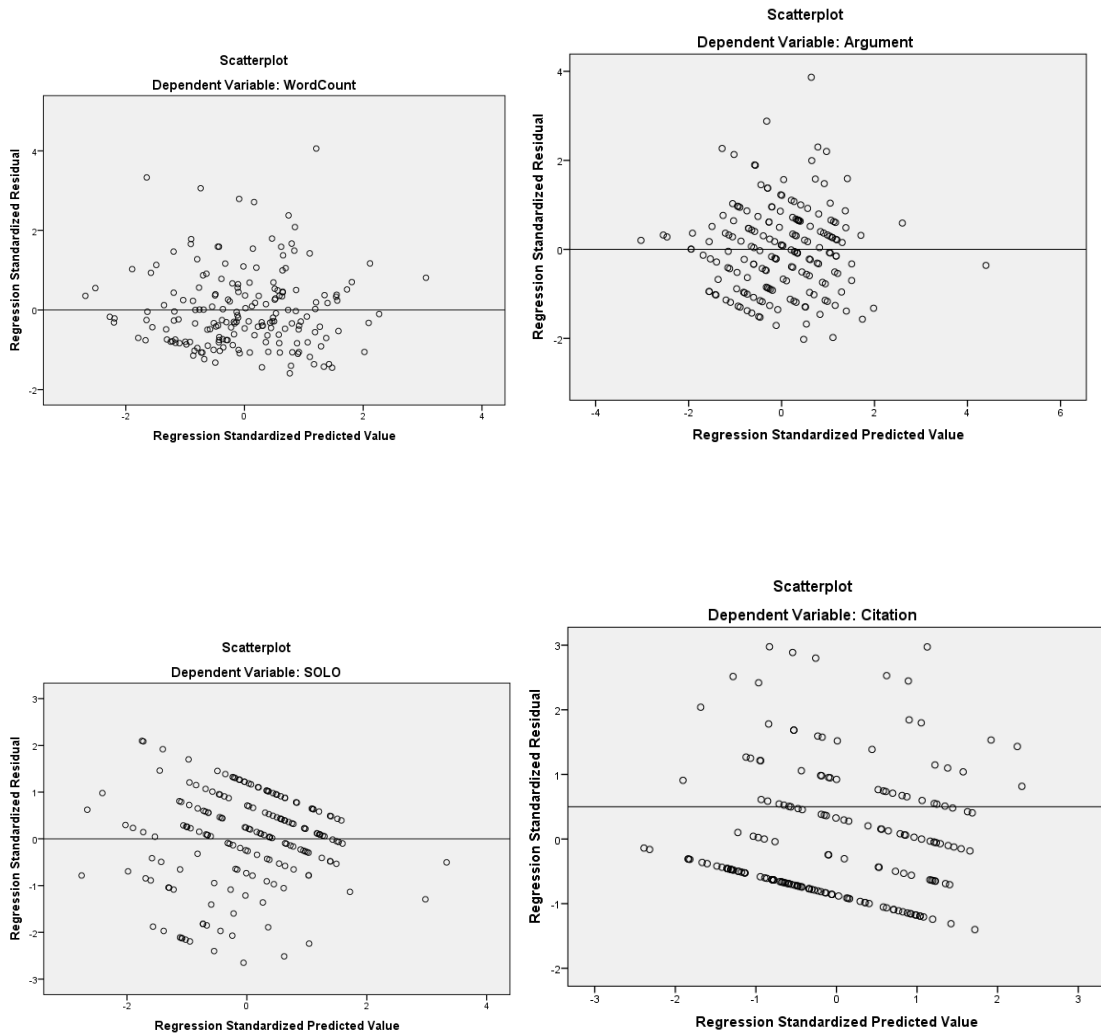
Residual plots using MSU behaviors to predict each outcome variable.



Residual plots using learner characteristics to predict each outcome variable.

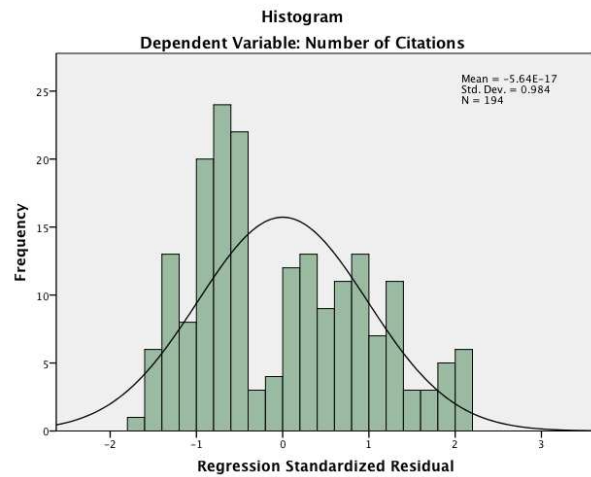
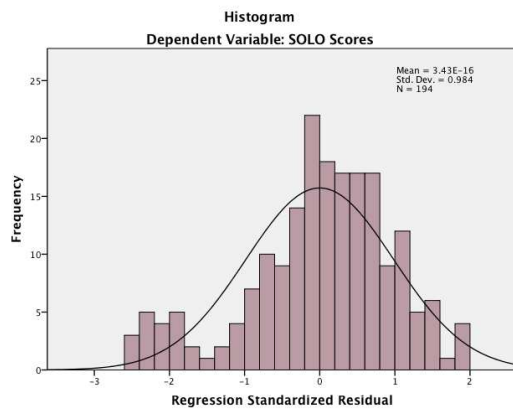
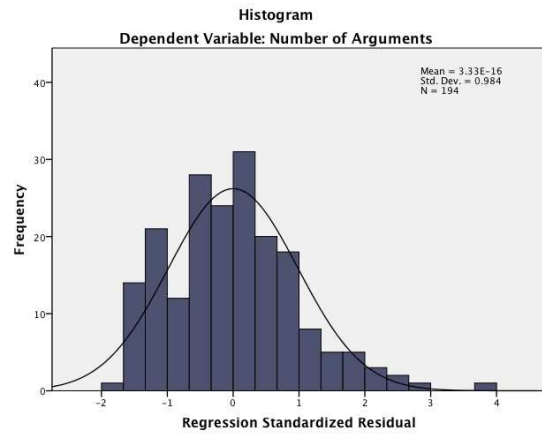
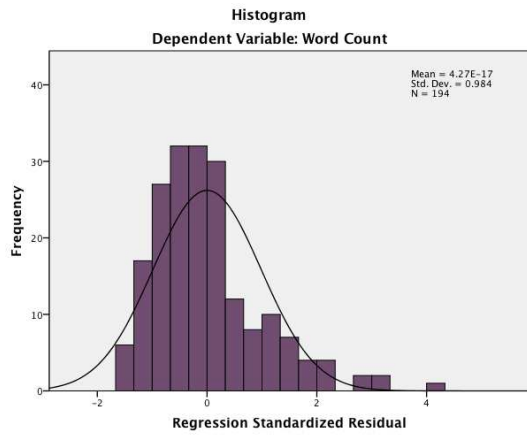


Residual plots using learner characteristics and MSU behaviors to predict each outcome variable.

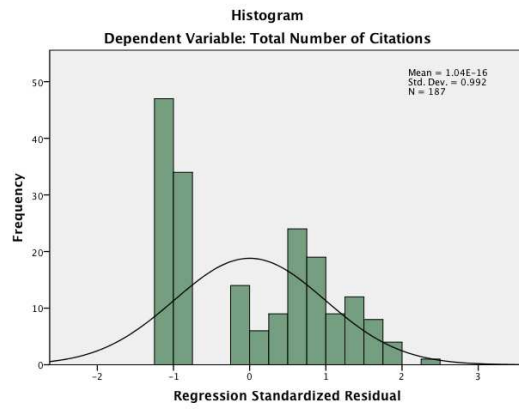
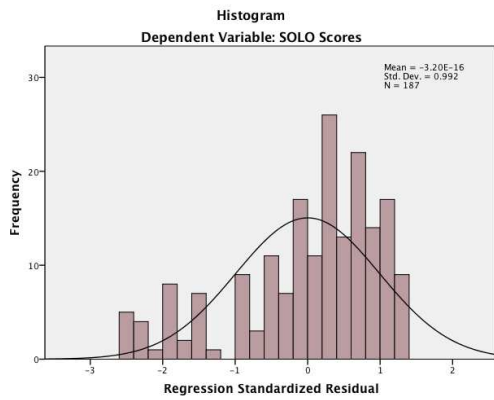
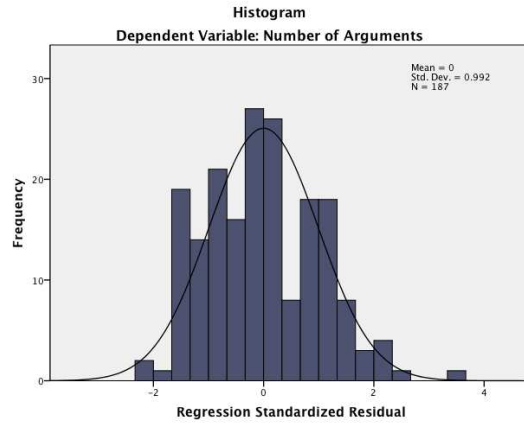
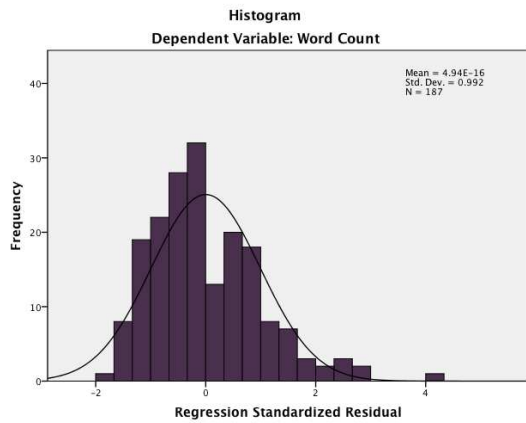


Appendix R: Histograms of Residuals for Predictive Models

Multiple Source Use Behaviors Predicting Open-Ended Response Quality

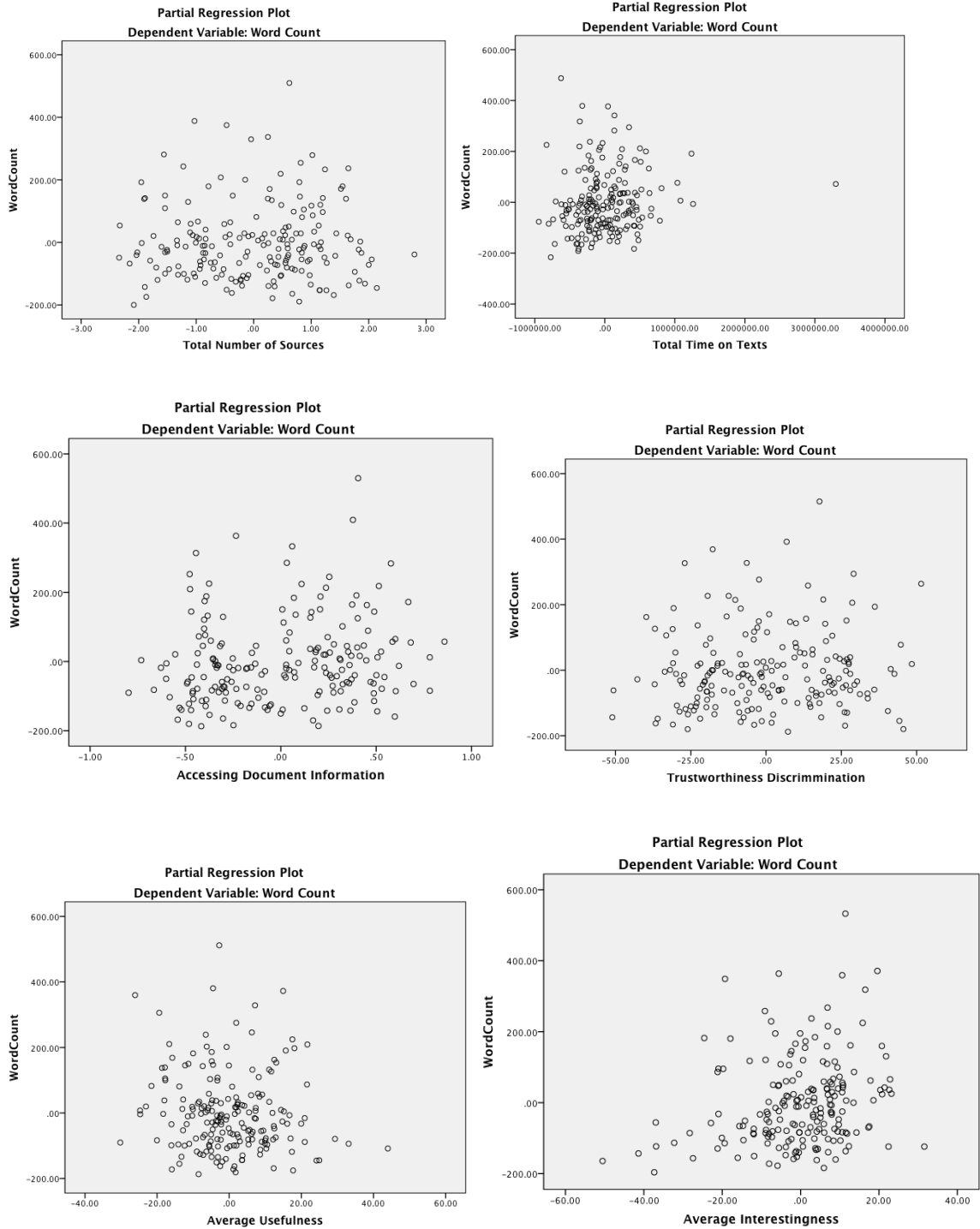


Learner Characteristics Predicting Open-Ended Response Quality

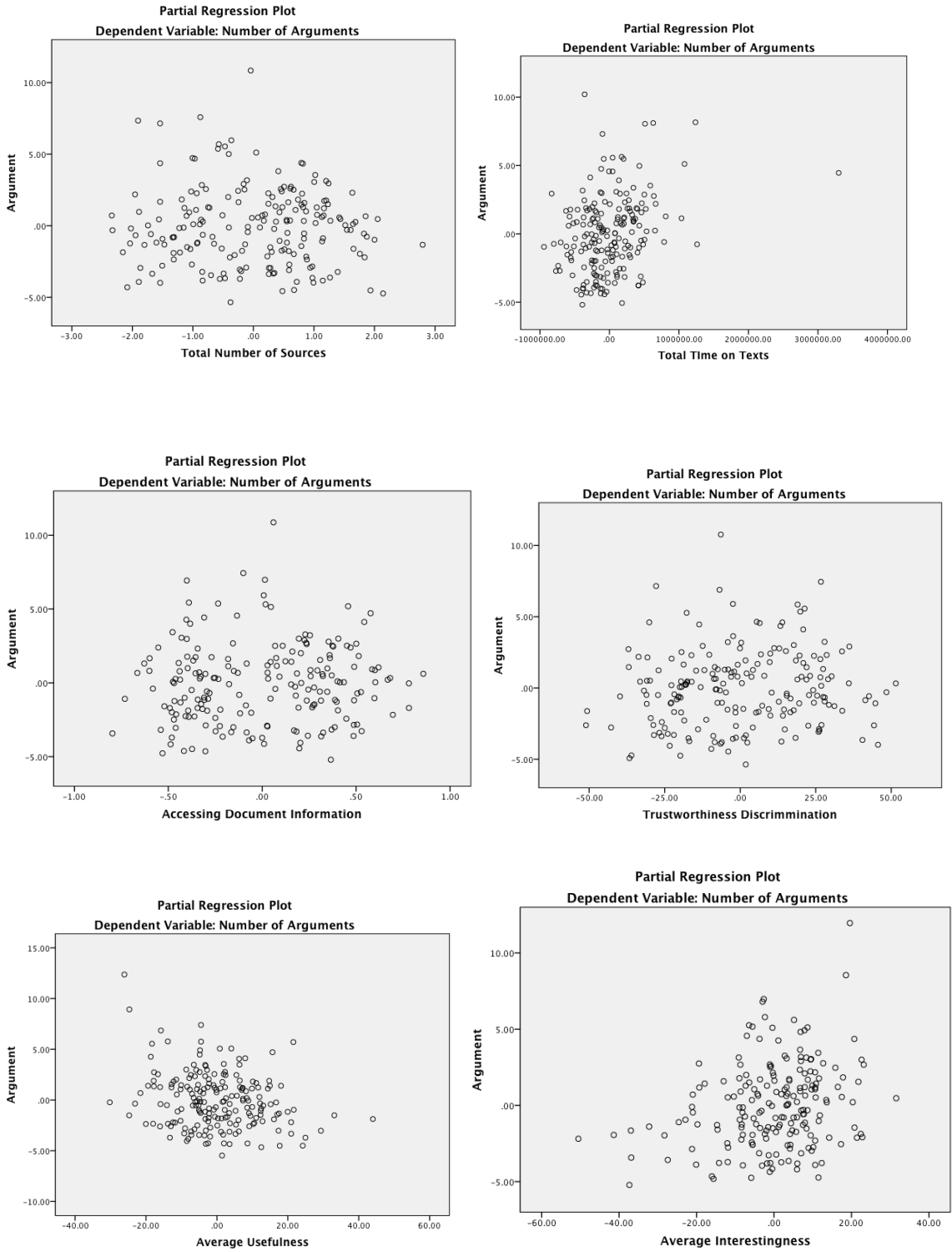


Appendix S: Partial Plots for Each Predictive Model

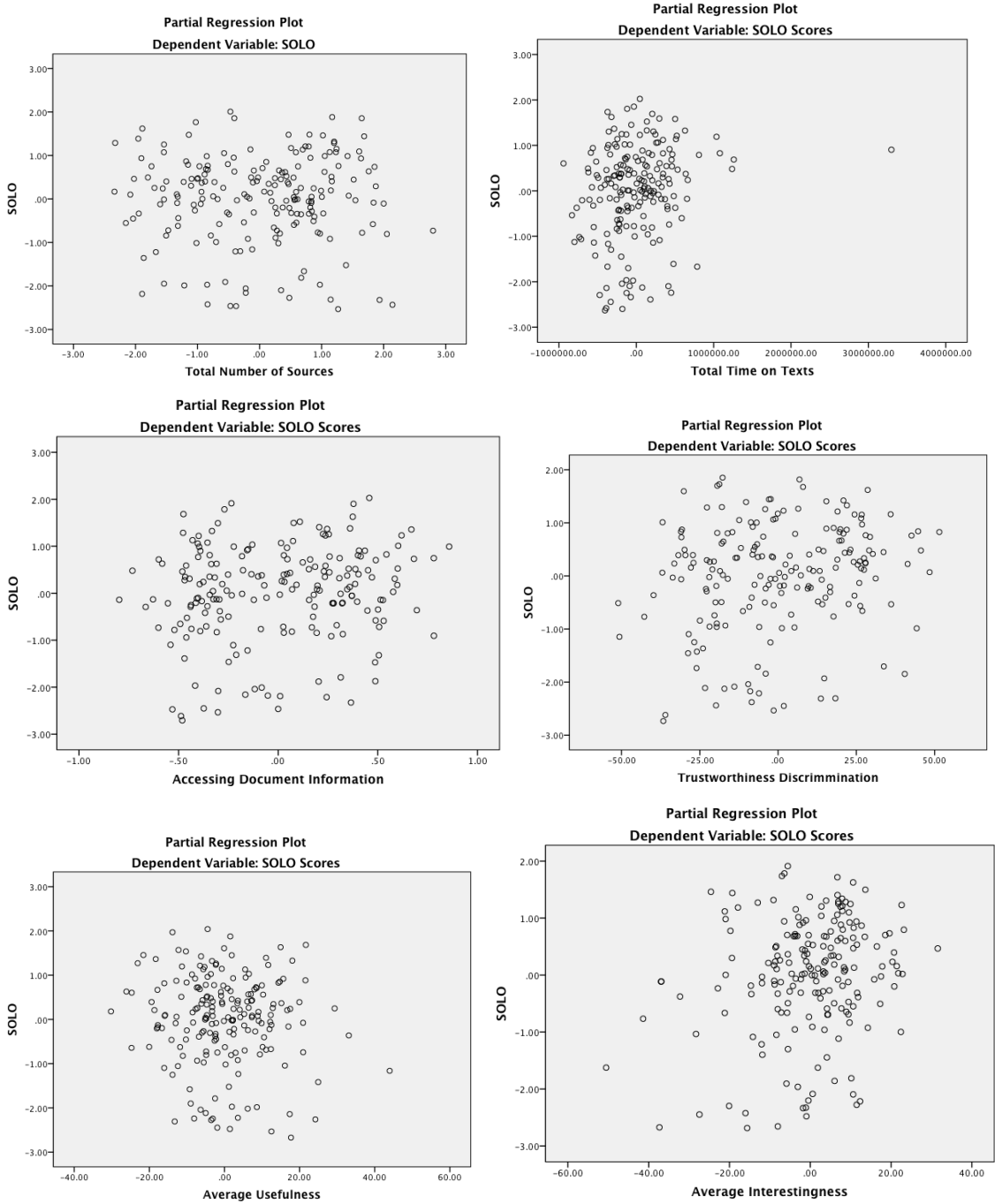
Partial plots of residuals of multiple source use behaviors predicting Word Count



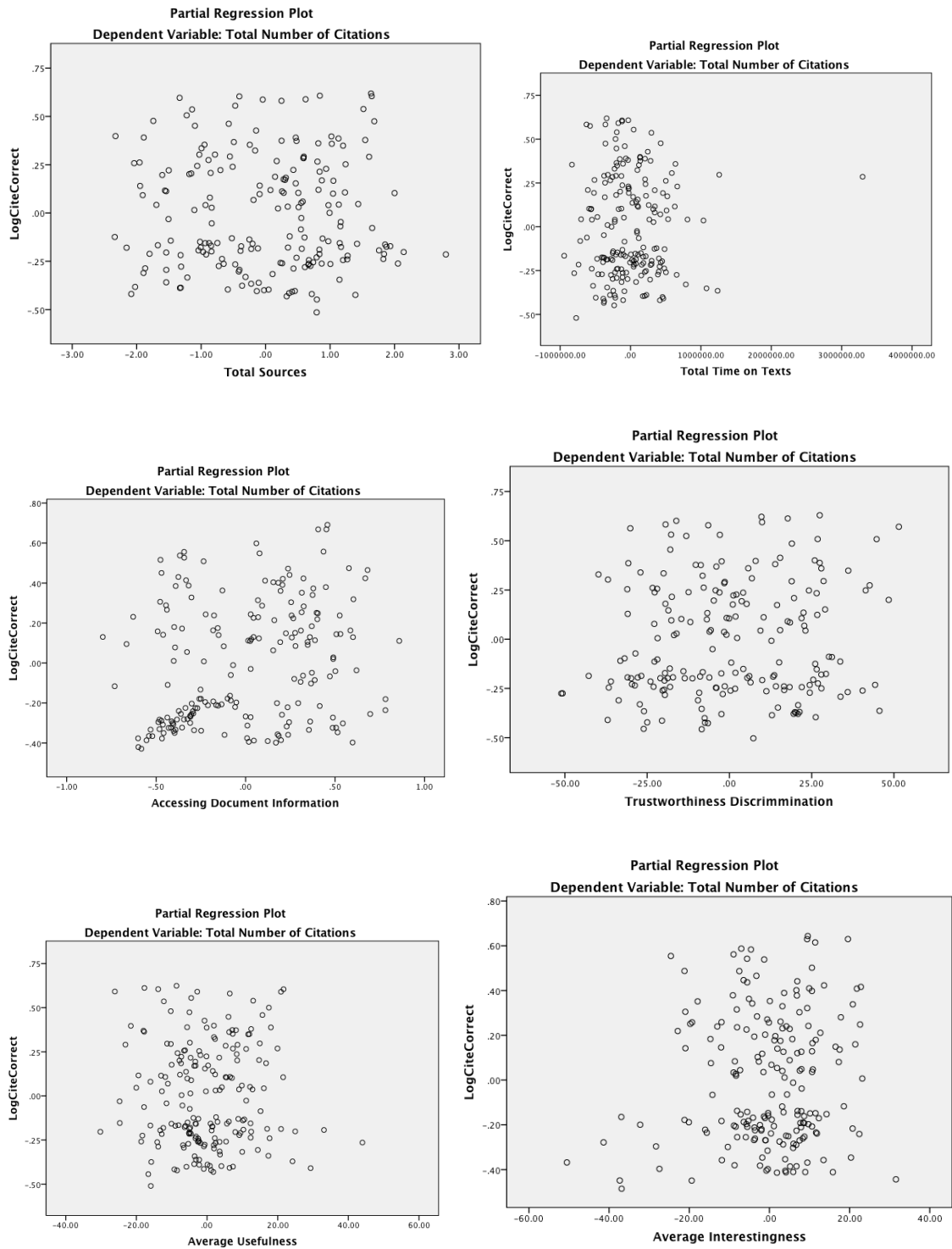
Partial plots of residuals of multiple source use behaviors predicting Arguments



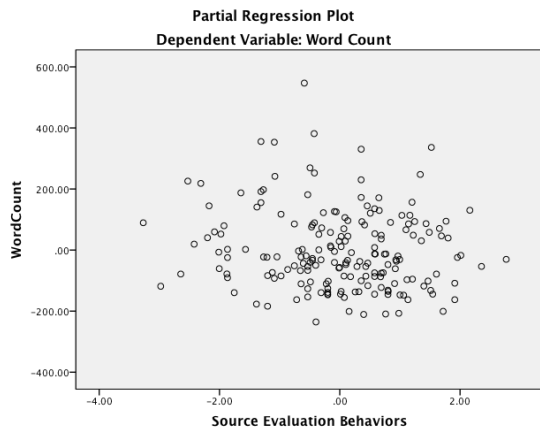
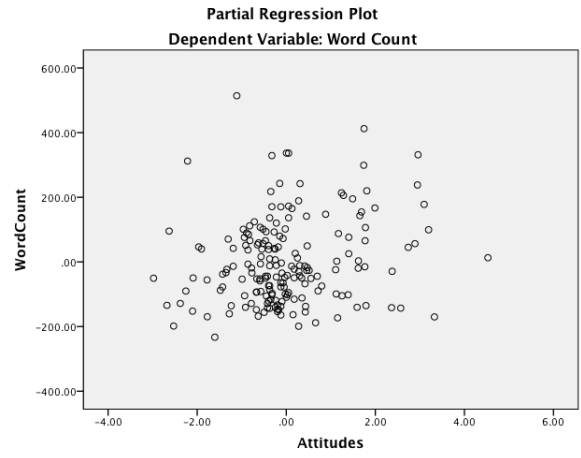
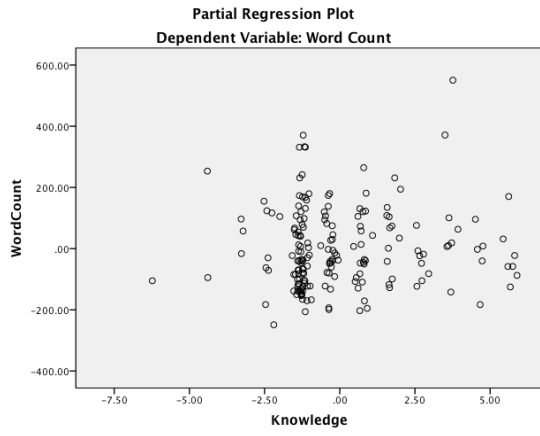
Partial plots of residuals of multiple source use behaviors predicting SOLO scores



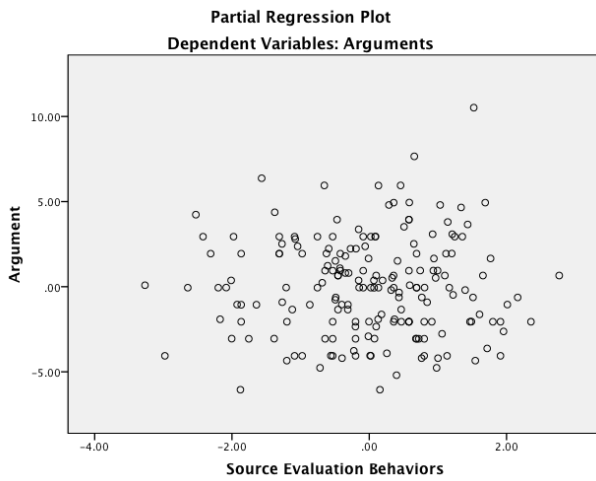
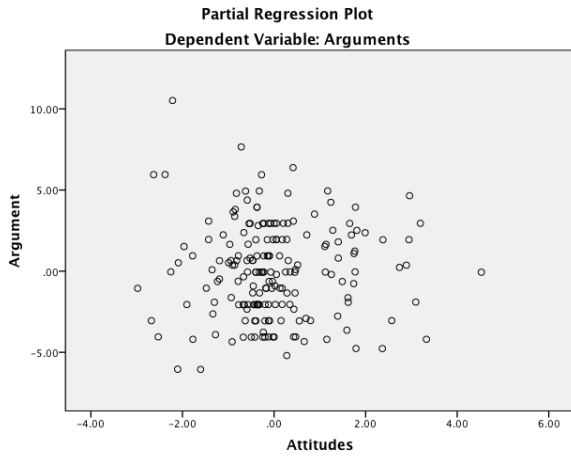
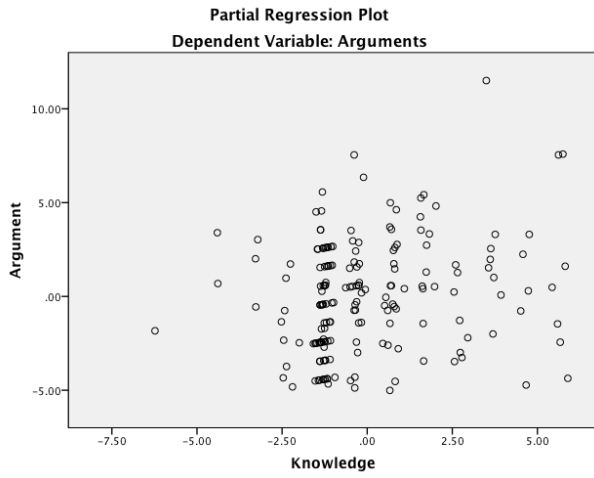
Partial plots of residuals of multiple source use behaviors predicting Citations



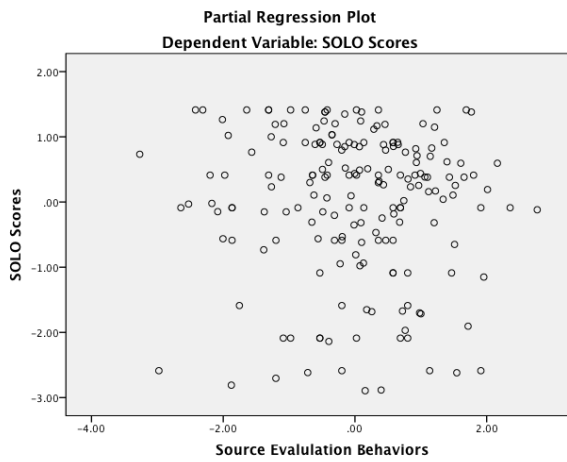
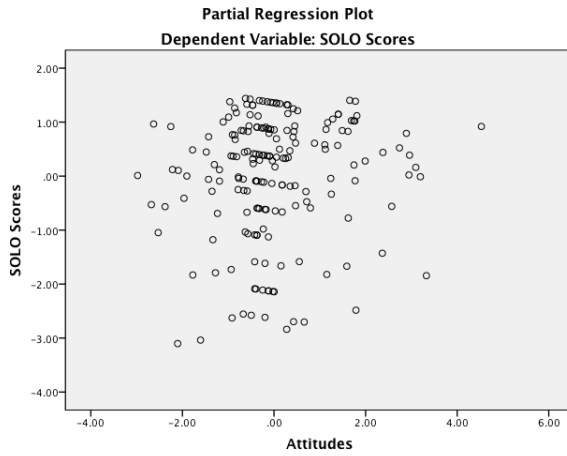
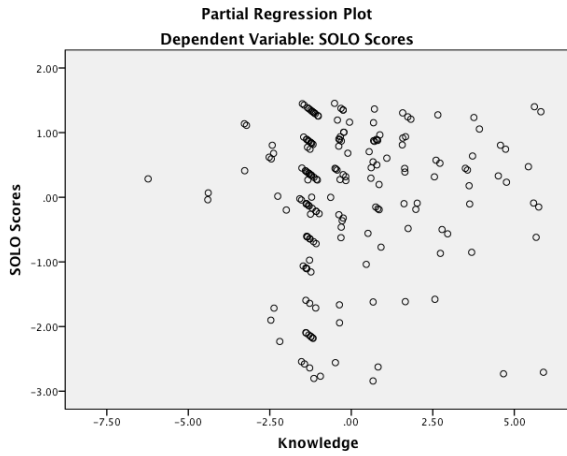
Partial plots of residuals of learner characteristics predicting Word Count



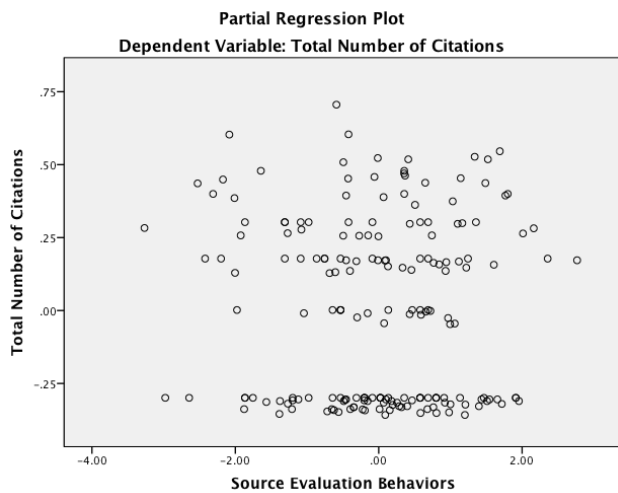
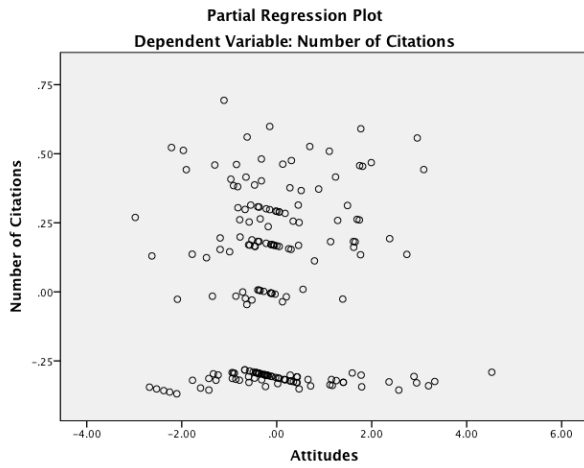
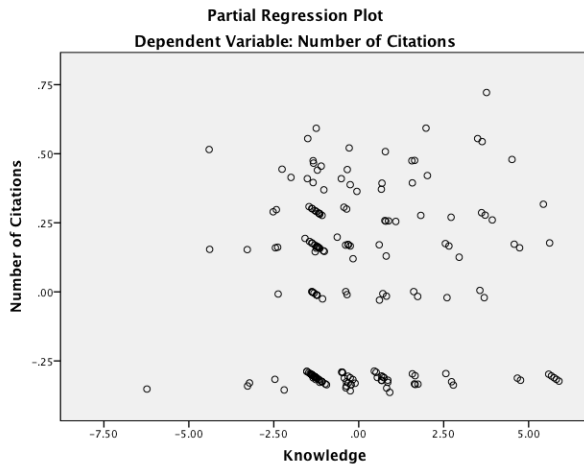
Partial plots of residuals of learner characteristics predicting Arguments



Partial plots of residuals of learner characteristics predicting SOLO Scores



Partial plots of residuals of learner characteristics predicting Citations



REFERENCES

- Afflerbach, P. (1990). The influence of prior knowledge on expert readers' main idea construction strategies. *Reading Research Quarterly*, 25(1), 31-46.
- Afflerbach, P., & Cho, B. (2009). Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. *Handbook of research on reading comprehension*, 69-90.
- Alexander, J. E., & Tate, M. A. (1999). *Web wisdom: How to evaluate and create information quality on the Web*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Alexander, P. A. (1997). Mapping the multidimensional nature of domain learning: The interplay of cognitive, motivational, and strategic forces. In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in Motivation and Achievement*, (Vol. 10, pp.213-250). Greenwich, CT: JAI Press Inc.
- Alexander, P. A. (2003). The development of expertise: the journey from acclimation to proficiency. *Educational Researcher*, 32(8), 10-14.
- Alexander, P. A., Kulikowich, J. M., & Jetton, T. L. (1994). The role of subject-matter knowledge and interest in the processing of linear and nonlinear texts. *Review of Educational Research*, 64(2), 201-252.
- Alexander, P. A., Kulikowich, J. M., & Schulze, S. K. (1994). How subject-matter knowledge affects recall and interest. *American Educational Research Journal*, 31(2), 313-337.
- Alexander, P. A., Murphy, P. K., Buehl, M. M., & Sperl, C. T. (1998). The influence of prior knowledge, beliefs, and interest in learning from persuasive text. In T.

- Shanahan & F. Rodriguez-Brown (Eds.), *Forty-Seventh Yearbook of the National Reading Conference* (pp. 167-181). Chicago, IL: National Reading Conference.
- Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: how researchers in learning and literacy talk about knowledge. *Review of Educational Research, 61*(3), 315-343.
- Amadiou, F., Tricot, A., & Mariné, C. (2009). Prior knowledge in learning from a non-linear electronic document: disorientation and coherence of the reading sequences. *Computers in Human Behavior, 25*(2), 381-388.
- Andre, T., & Windschitl, M. (2003). Interest, epistemological belief, and intentional conceptual change. In: G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional Conceptual Change* (pp.173-193). Mahwah, NJ: Erlbaum.
- Anmarkrud, Ø., McCrudden, M. T., Bråten, I., & Strømsø, H. I. (2013). Task-oriented reading of multiple documents: online comprehension processes and offline products. *Instructional Science, 41*(5), 873-894. doi: 10.1007/s11251-013-9263-8
- Ariasi, N., & Mason, L. (2011). Uncovering the effect of text structure in learning from a science text: an eye-tracking study. *Instructional Science, 39*(5), 581-601.
- Armstrong, C. L., & Collins, S. J. (2009). Reaching out: newspaper credibility among young adult readers. *Mass Communication and Society, 12*(1), 97-114.
- Attewell, P. (2001). Comment: The first and second digital divides. *Sociology of Education, 74*(3), 252-259.
- Azevedo, R. (2005). Computers as metacognitive tools for enhancing learning, *Educational Psychologist, 40*(4), 193-197.

- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173-1182.
- Beishuizen, J. J., & Stoutjesdijk, E. T. (1999). Study strategies in a computer assisted study environment. *Learning and Instruction*, *9*(3), 281-301.
- Bergkvist, L., & Rossiter, J. R. (2007). The predictive validity of multiple-item versus single-item measures of the same constructs. *Journal of Marketing Research*, *44*(2), 175-184.
- Bhabha, H. K. (1994). *The location of culture*. New York: Routledge.
- Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning*. New York: Academic Press
- Bishara, A. J., & Hittner, J. B. (2012). Testing the significance of a correlation with non-normal data: comparison of Pearson, Spearman, transformation, and resampling approaches. *Psychological Methods*, *17*(3), 399-417. doi:10.1037/a0028087
- Bland, J. M. & Altman, D. G. (1996a). Transforming data. *BMJ*, *312*, 770.
- Bland, J. M. & Altman, D. G. (1996b) The use of transformation when comparing two means. *BMJ*, *312*, 1153.
- Boscolo, P., & Mason, L. (2003). Topic knowledge, text coherence, and interest: how they interact in learning from instructional texts. *The Journal of Experimental Education*, *71*(2), 126-148.
- Boscolo, P., & Mason, L. (2003). Topic knowledge, text coherence, and interest: how they interact in learning from instructional texts. *The Journal of Experimental Education*, *71*(2), 126-148.

- Braasch, J. L., Bråten, I., Strømsø, H. I., & Anmarkrud, Ø. (2014). Incremental theories of intelligence predict multiple document comprehension. *Learning and Individual Differences, 31*, 11-20.
- Braasch, J. L., 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.
- Brand-Gruwel, S., Van Meeuwen, L., & Van Gog, T. (2008). The use of evaluation criteria when searching the WWW: an eye-tracking study. In: *Proceedings EARLI Special Interest Group Text And Graphics Comprehension 'Exploiting The Opportunities: Learning With Textual, Graphical, And Multimodal Representations'* (pp. 34-37).
- Brand-Gruwel, S., Wopereis, I., & Vermetten, Y. (2005). Information problem solving by experts and novices: analysis of a complex cognitive skill. *Computers in Human Behavior, 21*(3), 487-508.
- Brannon, L. A., Tagler, M. J., & Eagly, A. H. (2007). The moderating role of attitude strength in selective exposure to information. *Journal of Experimental Social Psychology, 43*(4), 611-617.
- Bråten, I., & Strømsø, H. I. (2003). A longitudinal think-aloud study of spontaneous strategic processing during the reading of multiple expository texts. *Reading and Writing, 16*(3), 195-218.

- Bråten, I., & Strømsø, H. I. (2006). Epistemological beliefs, interest, and gender as predictors of Internet-based learning activities. *Computers in Human Behavior, 22*(6), 1027-1042.
- Bråten, I., & Strømsø, H. I. (2011). Measuring strategic processing when students read multiple texts. *Metacognition and Learning, 6*(2), 111-130.
- Bråten, I., Anmarkrud, Ø., Brandmo, C., & Strømsø, H. I. (2014). Developing and testing a model of direct and indirect relationships between individual differences, processing, and multiple-text comprehension. *Learning and Instruction, 30*, 9-24.
- Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J. F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: toward an integrated model. *Educational Psychologist, 46*(1), 48-70.
- 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.
- Bråten, I., Strømsø, H. I., & Salmerón, L. (2011). Trust and mistrust when students read multiple information sources about climate change. *Learning and Instruction, 21*(2), 180-192.
- Brem, S. K., Russell, J., & Weems, L. (2001). Science on the Web: student evaluations of scientific arguments. *Discourse Processes, 32*(2-3), 191-213.
- Brenner, J. (2012). *Pew Internet: teens*. Report from the Pew Internet and American Life Project. Retrieved from: <http://pewinternet.org/Commentary/2012/April/Pew-Internet-Teens.aspx>

- Britt, M. A., & Aglinskias, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, 20(4), 485-522.
- Britt, M. A., & Gabrys, G. (2002). Implications of document-level literacy skills for Web site design. *Behavior Research Methods, Instruments, and Computers*, 34(2), 170-176.
- Britt, M. A., & Rouet, J. F. (2011). Research challenges in the use of multiple documents. *Information Design Journal*, 19(1), 62-67.
- Britt, M. A., Rouet, J. F., & Braasch, J. L. G. (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. 160-180). New York: Routledge, Taylor & Francis Group.
- Britt, M.A., Perfetti, C.A., Sandak, R., & Rouet, J. F., (1999). Content integration and source separation in learning from multiple texts. In Goldman, S.R., Graesser, A.C., & van den Broek, P. (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 73-90). Mahwah, NJ: Lawrence Erlbaum Associates.
- Buehl, M. M., Alexander, P. A., Murphy, P. K., & Sperl, C. T. (2001). Profiling persuasion: the role of beliefs, knowledge, and interest in the processing of persuasive texts that vary by argument structure. *Journal of Literacy Research*, 33(2), 269-301.
- Buehl, M. M., Alexander, P. A., Murphy, P. K., & Sperl, C. T. (2001). Profiling persuasion: the role of beliefs, knowledge, and interest in the processing of

- persuasive texts that vary by argument structure. *Journal of Literacy Research*, 33(2), 269-301.
- Burstein, J., Chodorow, M., & Leacock, C. (2004). Automated essay evaluation: the criterion online writing service. *Ai Magazine*, 25(3), 27-36.
- Burton, V. T., & Chadwick, S. A. (2000). Investigating the practices of student researchers: Patterns of use and criteria for use of Internet and library sources. *Computers and Composition*, 17(3), 309-328.
- Calkins, S., & Kelley, M. R. (2007). Evaluating Internet and scholarly sources across the disciplines: two case studies. *College Teaching*, 55(4), 151-156.
- Cerdán, R., & Vidal-Abarca, E. (2008). The effects of tasks on integrating information from multiple documents. *Journal of Educational Psychology*, 100(1), 209-222.
- Cerdán, R., Vidal-Abarca, E., Martínez, T., Gilabert, R., & Gil, L. (2009). Impact of question-answering tasks on search processes and reading comprehension. *Learning and Instruction*, 19(1), 13-27.
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752-766
- Churchill Jr., G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 16(1), 64-73.
- Ciolek, T. M. (1996). The six quests for the electronic grail: current approaches to information quality in WWW resources. *Revue Informatique et Statistique dans les Sciences Humaines*, 32, 1-4.

- Cleveland, W. S. (1984). Graphic methods for data presentation: full scale breaks, dot charts, and multibased logging. *The American Statistician*, 38(4), 270-280.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2002) *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
- Coiro, J. (2003a) Reading comprehension on the Internet: expanding our understanding of reading comprehension to encompass new literacies. *The Reading Teacher*, 56(5), 458-464.
- Coiro, J. (2003b). Rethinking comprehension strategies to better prepare students for critically evaluating content on the Internet. *The New England Reading Association Journal*, 39(2), 29-34.
- Common Core State Standards Initiative (2010). *Common Core State Standards for English Language Arts and Literacy in History/Social Studies, Science, and Technical Subjects*. Washington, DC: Council of Chief State School Officers and the National Governors Association Center for Best Practices.
- Cook, R. D., & Weisberg, S. (1983). Diagnostics for heteroscedasticity in regression. *Biometrika*, 70(1), 1-10.
- Cooper, J. (2006). The digital divide: the special case of gender. *Journal of Computer Assisted Learning*, 22(5), 320-334.
- Cottrell, J. R. (2001). Teaching students to evaluate web sources more critically: Implications from a faculty workshop. *College and Research Libraries News*, 62(2), 141-143.

- Dillon, A., & Gushrowski, B. A. (2000). Genres and the Web: is the personal home page the first uniquely digital genre? *Journal of the American Society for Information Science*, *51*(2), 202-205.
- Dole, J. A., & Sinatra, G. M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist*, *33*(2-3), 109-128.
- Drachman, D. (2012). Adjustments for multiple comparisons. *Journal of Clinical Research Best Practices*, *8*(7), 1-3.
- Dreher, M. J., & Guthrie, J. T. (1990). Cognitive processes in textbook chapter search tasks. *Reading Research Quarterly*, *25*(4), 323-339.
- Duggan, G. B., & Payne, S. J. (2011, May). Skim reading by satisficing: evidence from eye tracking. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1141-1150).
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.
- Eagly, A. H., Chen, S., Chaiken, S., & Shaw-Barnes, K. (1999). The impact of attitudes on memory: an affair to remember. *Psychological Bulletin*, *125*(1), 64-89.
- Eagly, A. H., Kulesa, P., Chen, S., & Chaiken, S. (2001). Do attitudes affect memory?: tests of the congeniality hypothesis. *Current Directions in Psychological Science*, *10*(1), 5-9.
- Edwards, K., & Smith, E. E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology*, *71*(1), 5-24.
- Eliopoulos, D., & Gotlieb, C. (2003). Evaluating web search results rankings. *Online*, *27*, 42-48.

Environmental Forensics 2,359-362.

Ericsson, K. A., & Smith, J. (Eds.). (1991). *Toward a general theory of expertise:*

Prospects and limits. New York: Cambridge University Press.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149-1160.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.

Feise, R. J. (2002). Do multiple outcome measures require p-value adjustment? *BMC Medical Research Methodology*, 2(1), 8-12.

Ferguson, L. E., Bråten, I., & Strømsø, H. I. (2012). Epistemic cognition when students read multiple documents containing conflicting scientific evidence: a think-aloud study. *Learning and Instruction*, 22(2), 103-120.

Ferro, D., & Swedin, E. (2009). Computer fiction: “a logic named Joe.” In J. Impagliazzo, T. Järvi, & P. Paju (Eds.), *History of Nordic Computing* 2, (pp.84-94) Berlin: Springer.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. (4th ed.). London: Sage.

Fincher-Kiefer, R. (1992). The role of prior knowledge in inferential processing. *Journal of Research in Reading*, 15(1), 12-27.

Fischer, P., & Greitemeyer, T. (2010). A new look at selective-exposure effects an integrative model. *Current Directions in Psychological Science*, 19(6), 384-389.

- Fischer, P., Jonas, E., Frey, D., & Schulz-Hardt, S. (2005). Selective exposure to information: the impact of information limits. *European Journal of Social Psychology, 35*(4), 469-492.
- Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of Internet information credibility. *Journalism and Mass Communication Quarterly, 77*(3), 515-540.
- Flowerdew, L. (2000). Using a genre-based framework to teach organizational structure in academic writing. *ELT journal, 54*(4), 369-378.
- Foltz, P. W. (1996). Comprehension, coherence, and strategies in hypertext and linear text. In J. Levonen, J. F. Rouet, A. Dillon, & R. Spiro (Eds.), *Hypertext and Cognition*, (pp. 109–136). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Fowler, R. (1987). Power and robustness in product-moment correlation. *Applied Psychological Measurement, 11*(4), 419-428.
- Garner, R., Alexander, P. A., Gillingham, M. G., Kulikowich, J. M., & Brown, R. (1991). Interest and learning from text. *American Educational Research Journal, 28*(3), 643-659.
- Garner, R., Brown, R., Sanders, S., & Menke, D. J. (1992). Seductive details' and learning from text. In K. A. Renning, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 239-254).
- Garner, R., Gillingham, M. G., & White, C. S. (1989). Effects of 'seductive details' on macroprocessing and microprocessing in adults and children. *Cognition and Instruction, 6*(1), 41-57.
- Gauthier, T. (2001). Detecting trends using Spearman's rank correlation coefficient.

- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.
- Gerjets, P., Kammerer, Y., & Werner, B. (2011). Measuring spontaneous and instructed evaluation processes during Web search: integrating concurrent thinking-aloud protocols and eye-tracking data. *Learning and Instruction, 21*(2), 220-231.
- Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H. I. (2010a). Summary versus argument tasks when working with multiple documents: which is better for whom? *Contemporary Educational Psychology, 35*(3), 157-173.
- Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H. I. (2010b). Understanding and integrating multiple science texts: summary tasks are sometimes better than argument tasks. *Reading Psychology, 31*(1), 30-68.
- Goldman, S. R., & Scardamalia, M. (2013a). Managing, understanding, applying, and creating knowledge in the information age: next-generation challenges and opportunities. *Cognition and Instruction, 31*(2), 255-269.
- Goldman, S. R., & Scardamalia, M. (2013b). Preface for the special issue multiple document comprehension. *Cognition and Instruction, 31*(2), 121-121.
- Goldman, S. R., Braasch, J. L., 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.
- Goodwin, J. (2011). Why we were surprised (again) by the Arab Spring. *Swiss Political Science Review, 17*(4), 452-456.
- Graf, J., & Aday, S. (2008). Selective attention to online political information. *Journal of Broadcasting & Electronic Media, 52*(1), 86-100.

- Graham, L., & Metaxas, P. T. (2003). "Of course it's true; I saw it on the Internet!": critical thinking in the Internet era. *Communications of the Association for Computing Machinery*, 46(5), 70-75.
- Grimes, D. J., & Boening, C. H. (2001). Worries with the Web: a look at student use of Web resources. *College and Research Libraries*, 62(1), 11-22.
- Guthrie, J. T. (1988). Locating information in documents: examination of a cognitive model. *Reading Research Quarterly*, 23(2), 178-199.
- Guthrie, J. T., & Kirsch, I. S. (1987). Distinctions between reading comprehension and locating information in text. *Journal of Educational Psychology*, 79(3), 220-227.
- Guthrie, J. T., Wigfield, A., Metsala, J. L., & Cox, K. E. (1999). Motivational and cognitive predictors of text comprehension and reading amount. *Scientific Studies of Reading*, 3(3), 231-256.
- Hargittai, E., Fullerton, L., Menchen-Trevino, E., & Thomas, K. Y. (2010). Trust online: young adults' evaluation of web content. *International Journal of Communication*, 4(1), 468-494.
- Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: a theory of cognitive interest in science learning. *Journal of Educational Psychology*, 90(3), 414-434.
- Hartzler-Miller, C. (2001). Making sense of "best practice" in teaching history. *Theory & Research in Social Education*, 29(4), 672-695.
- Head, A. J., & Eisenberg, M. B. (2010). How today's college students use Wikipedia for course-related research. *First Monday*, 15(3). Retrieved from: <http://uncommonculture.org/ojs/index.php/fm/article/view/2830>

- 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.
- Hicks, D., Doolittle, P., & Lee, J. K. (2004). Social studies teachers' use of classroom-based and web-based historical primary sources. *Theory & Research in Social Education*, 32(2), 213-247.
- Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research*, 60(4), 549-571.
- Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations. *Educational Psychology Review*, 13(3), 191-209.
- Hidi, S. (2006). Interest: a unique motivational variable. *Educational Research Review*, 1(2), 69-82.
- Hidi, S., & Anderson, V. (1992). Situational interest and its impact on reading and expository writing. In K. A. Renning, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 215-238). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hidi, S., & Baird, W. (1986). Interestingness—a neglected variable in discourse processing. *Cognitive Science*, 10(2), 179-194.
- Hidi, S., & Baird, W. (1988). Strategies for increasing text-based interest and students' recall of expository texts. *Reading Research Quarterly*, 23(4), 465-483.
- Hidi, S., Renninger, K. A., & Krapp, A. (2004). Interest, a motivational variable that combines affective and cognitive functioning. In D. Y. Dai & R. J. Sternberg (Eds.) *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development*, (pp. 89-115.) Mahwah, NJ: Lawrence Erlbaum.

- Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88-140.
- Hofer, B. K., & Pintrich, P. R. (Eds.). (2004). *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hofman, R., & Van Oostendorp, H. (1999). Cognitive effects of a structural overview in a hypertext. *British Journal of Educational Technology*, 30(2), 129-140.
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion: Psychological studies of opinion change*. New Haven, CT: Yale University Press.
- Hynd, C., Holschuh, J. P., & Hubbard, B. P. (2004). Thinking like a historian: college students' reading of multiple historical documents. *Journal of Literacy Research*, 36(2), 141-176.
- Johnson, T. J., & Kaye, B. K. (1998). Cruising is believing?: comparing Internet and traditional sources on media credibility measures. *Journalism & Mass Communication Quarterly*, 75(2), 325-340.
- Judd, C. M., & Kulik, J. A. (1980). Schematic effects of social attitudes on information processing and recall. *Journal of Personality and Social Psychology*, 38(4), 569-578.
- Kammerer, Y., & Gerjets, P. (2010). How the interface design influences users' spontaneous trustworthiness evaluations of web search results: comparing a list and a grid interface. Proceedings from ETRA 2010: *Symposium on Eye-Tracking Research and Applications* (pp. 299-306). Association for Computing Machinery.

- Kammerer, Y., & Gerjets, P. (2012a). Effects of search interface and Internet-specific epistemic beliefs on source evaluations during Web search for medical information: an eye-tracking study. *Behavior & Information Technology*, *31*(1), 83-97.
- Kammerer, Y., & Gerjets, P. (2012b). How search engine users evaluate and select Web search results: the impact of the search engine interface on credibility assessments. *Library and Information Science*, *4*, 251-279. doi: [10.1108/S1876-0562\(2012\)002012a012](https://doi.org/10.1108/S1876-0562(2012)002012a012)
- Kammerer, Y., Bråten, I., Gerjets, P., & Strømsø, H. I. (2013). The role of Internet-specific epistemic beliefs in laypersons' source evaluations and decisions during Web search on a medical issue. *Computers in Human Behavior*, *29*(3), 1193-1203.
- Kammerer, Y., Wollny, E., Gerjets, P., & Scheiter, K. (2009). How authority-related epistemological beliefs and salience of source information influence the evaluation of Web search results—an eye tracking study. In: *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2158-2163).
- Kelley, K., & Maxwell, S. E. (2010). Multiple regression. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences*. New York: Taylor & Francis.
- Kelly, D., & Belkin, N. J. (2001, September). Reading time, scrolling and interaction: exploring implicit sources of user preferences for relevance feedback. In *Proceedings of the 24th annual international ACM SIGIR conference on*

- Research and development in information retrieval* (pp. 408-409). New York: Springer-Verlag.
- Keppel, G., & Zedek, S. (1989). *Data analysis for research designs: Analysis of variance and multiple regression/correlation approaches*. New York: W. H. Freeman.
- Kerstetter, D., & Cho, M. (2004). Prior knowledge, credibility and information search. *Annals of Tourism Research*, 31(4), 961-985.
- Keselman, H. J., Algina, J., Lix, L. M., Wilcox, R. R., & Deering, K. N. (2008). A generally robust approach for testing hypotheses and setting confidence intervals for effect sizes. *Psychological Methods*, 13(2), 110-129.
- Khondker, H. H. (2011). Role of the new media in the Arab Spring. *Globalizations*, 8(5), 675-679.
- Kienhues, D., Stadler, M., & Bromme, R. (2011). Dealing with conflicting or consistent medical information on the Web: when expert information breeds laypersons' doubts about experts. *Learning and Instruction*, 21(2), 193-204.
- Kiili, C., Laurinen, L., & Marttunen, M. (2008). Students evaluating Internet sources: from versatile evaluators to uncritical readers. *Journal of Educational Computing Research*, 39(1), 75-95.
- Kim, H. J. J., & Millis, K. (2006). The influence of sourcing and relatedness on event integration. *Discourse Processes*, 41, 51-65. doi: 10.1207/s15326950dp4101_4
- King, P. M., & Kitchener, K. S. (2004). Reflective judgment: theory and research on the development of epistemic assumptions through adulthood. *Educational Psychologist*, 39(1), 5-18.

- Kintsch, W. (1980). Learning from text, levels of comprehension, or: why anyone would read a story anyway. *Poetics*, 9, 87–89.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychological Review*, 95(2), 163-182.
- Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, 49(4), 294-303.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363-394.
- Kintsch, W., Kozminsky, E., Streby, W. J., McKoon, G., & Keenan, J. M. (1975). Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior*, 14(2), 196-214.
- Kitchener, K. S., & King, P. M. (1981). Reflective judgment: concepts of justification and their relationship to age and education. *Journal of Applied Developmental Psychology*, 2(2), 89-116.
- Knobloch-Westerwick, S., & Meng, J. (2011). Reinforcement of the political self through selective exposure to political messages. *Journal of Communication*, 61(2), 349-368.
- Kobayashi, K. (2010). Strategic use of multiple texts for the evaluation of arguments. *Reading Psychology*, 31(2), 121-149.
- Krapp, A. (1999). Interest, motivation and learning: an educational-psychological perspective. *European Journal of Psychology of Education*, 14(1), 23-40.

- Krosnick, J. A., & Petty, R. E. (1995). Attitude strength: an overview. In: R. E. Petty & J. A. Krosnick (Eds.), *Attitude Strength: Antecedents and consequences*, (pp. 1-24). Mahwah, NJ: Lawrence Erlbaum Associates.
- Kuhn, D. (1991). *The skills of argument*. New York: Cambridge University Press
- Kuiper, E., Volman, M., & Terwel, J. (2005). The Web as an information resource in K-12 education: strategies for supporting students in searching and processing information. *Review of Educational Research*, 75(3), 285-328.
- Kunkel, L. R., Weaver, S. M., & Cook, K. N. (1996). What do they know?: an assessment of undergraduate library skills. *The Journal of Academic Librarianship*, 22(6), 430-434.
- Kushin, M. J., & Yamamoto, M. (2010). Did social media really matter? college students' use of online media and political decision making in the 2008 election. *Mass Communication and Society*, 13(5), 608-630.
- Langer, J. A. (1982). Facilitating text processing: The elaboration of prior knowledge. In J. A. Langer & M. Smith-Burke (Eds.), *Reader meets author/Bridging the gap: A psycholinguistic and sociolinguistic perspective*. Newark, DE: International Reading Association.
- Lankshear, C., & Knobel, M. (2011). *New literacies: Everyday practices and social learning*. New York: McGraw-Hill International.
- Lattin, J. M., Carroll, J. D., & Green, P. E. (2003). *Analyzing multivariate data* (pp. 38-40). Pacific Grove, CA: Thomson Brooks/Cole.

- Le Bigot, L., & Rouet, J. F. (2007). The impact of presentation format, task assignment, and prior knowledge on students' comprehension of multiple online documents. *Journal of Literacy Research, 39*(4), 445-470.
- Leinster, M. (2005). *A Logic Named Joe*. Riverdale, NY: Baen Publishing Enterprises.
- Leu, D. J., Kinzer, C. K., Coiro, J. L., & Cammack, D. W. (2004). Toward a theory of new literacies emerging from the Internet and other information and communication technologies. *Theoretical Models and Processes of Reading, 5*, 1570-1613.
- Leu, D. J., Kinzer, C. K., Coiro, J. L., & Cammack, D. W. (2004). Toward a theory of new literacies emerging from the Internet and other information and communication technologies. *Theoretical Models and Processes of Reading, 5*(1), 1570-1613.
- Levine, J. M., & Murphy, G. (1943). The learning and forgetting of controversial material. *The Journal of Abnormal and Social Psychology, 38*(4), 507.
- Lim, S. (2009). How and why do college students use Wikipedia? *Journal of the American Society for Information Science and Technology, 60*(11), 2189-2202.
- Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*(11), 2098-2109.
- Lorenzen, M. (2001). The land of confusion?: high school students and their use of the World Wide Web for research. *Research Strategies, 18*(2), 151-163.

- Lotan, G., Graeff, E., Ananny, M., Gaffney, D., & Pearce, I. (2011). The revolutions were tweeted: information flows during the 2011 Tunisian and Egyptian revolutions. *International Journal of Communication*, 5, 1375-1405.
- Lycke, K. H., Strømsø, H. I., & Grøttum, P. (2002). *PBL goes ICT: problem-based learning in face-to-face and distributed groups in medical education at the University of Oslo*. Oslo, Norway: Institute for Educational Research, University of Oslo.
- Macedo-Rouet, M., Braasch, J. L., Britt, M. A., & Rouet, J. F. (2013). Teaching fourth and fifth graders to evaluate information sources during text comprehension. *Cognition and Instruction*, 31(2), 204-226.
- Maier, J., & Richter, T. (2013). Text belief consistency effects in the comprehension of multiple texts with conflicting information. *Cognition and Instruction*, 31(2), 151-175.
- Mason, L., & Boscolo, P. (2004). Role of epistemological understanding and interest in interpreting a controversy and in topic-specific belief change. *Contemporary Educational Psychology*, 29(2), 103-128.
- Mason, L., Ariasi, N., & Boldrin, A. (2011). Epistemic beliefs in action: spontaneous reflections about knowledge and knowing during online information searching and their influence on learning. *Learning and Instruction*, 21(1), 137-151.
- Mason, L., Boldrin, A., & Ariasi, N. (2010a). Epistemic metacognition in context: evaluating and learning online information. *Metacognition and Learning*, 5(1), 67-90.

- Mason, L., Boldrin, A., & Ariasi, N. (2010b). Searching the Web to learn about a controversial topic: are students epistemically active? *Instructional Science*, 38(6), 607-633.
- McCrudden, M. T., & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review*, 19(2), 113-139.
- McCrudden, M. T., & Sparks, P. C. (2014). Exploring the effect of task instructions on topic beliefs and topic belief justifications: A mixed methods study. *Contemporary Educational Psychology*, 39(1), 1-11.
- McCrudden, M. T., Magliano, J. P., & Schraw, G. (2010). Exploring how relevance instructions affect personal reading intentions, reading goals and text processing: a mixed methods study. *Contemporary Educational Psychology*, 35(4), 229-241.
- McCrudden, M. T., Schraw, G., & Hartley, K. (2006). The effect of general relevance instructions on shallow and deeper learning and reading time. *The Journal of experimental education*, 74(4), 291-310.
- McDonald, S., & Stevenson, R. J. (1998). Navigation in hyperspace: an evaluation of the effects of navigational tools and subject matter expertise on browsing and information retrieval in hypertext. *Interacting with Computers*, 10(2), 129-142.
- McInerney, C. R., & Bird, N. J. (2005). Assessing website quality in context: retrieving information about genetically modified food on the Web. *Information Research*, 10(2). Retrieved from:
http://libres.uncg.edu/ir/uncg/f/N_Bird%20Assessing_2005.pdf

- McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*(1), 1-43.
- Metzger, M. (2005). Understanding how Internet users make sense of credibility: A review of the state of our knowledge and recommendations for theory, policy, and practice. In proceedings from American Library Association: *Symposium on Internet Credibility and the User*.
- Metzger, M. J. (2007). Making sense of credibility on the Web: models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology, 58*(13), 2078-2091.
- Metzger, M. J., Flanagin, A. J., & Zwarun, L. (2003). College student Web use, perceptions of information credibility, and verification behavior. *Computers and Education, 41*(3), 271-290.
- Miller, L. M. S. (2003). The effects of age and domain knowledge on text processing. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 58*(4), 217-223.
- Moje, E. B., Ciechanowski, K. M., Kramer, K., Ellis, L., Carrillo, R., & Collazo, T. (2004). Working toward third space in content area literacy: an examination of everyday funds of knowledge and discourse. *Reading Research Quarterly, 39*(1), 38-70.
- Möller, J. & Müller-Kalthoff, T. (2000) Learning with hypertext: the impact of navigational aids and prior knowledge. *German Journal of Educational Psychology, 14*(2/3), 116-123.

- Morita, M., & Shinoda, Y. (1994, August). Information filtering based on user behavior analysis and best match text retrieval. In *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 272-281). New York: Springer-Verlag.
- Mosenthal, P. B. (1996). Understanding the strategies of document literacy and their conditions of use. *Journal of Educational Psychology*, 88(2), 314-322.
- Munro, G. D., & Ditto, P. H. (1997). Biased assimilation, attitude polarization, and affect in reactions to stereotype-relevant scientific information. *Personality and Social Psychology Bulletin*, 23(6), 636-653.
- Murphy, P. K., & Mason, L. (2006). Changing knowledge and beliefs. In P. A. Alexandre & P. H. Winne (Eds.), *Handbook of Educational Psychology*, (pp. 305-324). Mahwah, NJ: Lawrence Erlbaum Associates.
- Nan, X. (2009). The influence of source credibility on attitude certainty: exploring the moderating effects of timing of source identification and individual need for cognition. *Psychology & Marketing*, 26(4), 321-332.
- Nickerson, R. S. (1998). Confirmation bias: a ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175-220.
- Nimon, K. F. (2012). Statistical assumptions of substantive analyses across the general linear model: a mini-review. *Frontiers in psychology*, 3,1-5.
- Northwestern University, (Modified: 1997). *PROPHET StateGuide*. Retrieved from: http://www.basic.northwestern.edu/statguidefiles/linreg_ass_viol.html
- O'Keefe, D. J. (2003). Colloquy: should family-wise alpha be adjusted?. *Human Communication Research*, 29(3), 431-445.

- Osborne, J. (2002). Notes on the use of data transformations. *Practical Assessment, Research & Evaluation, 8*(6), 1-8. Retrieved from:
<http://pareonline.net/getvn.asp?v=8&n=6>.
- Osborne, J. W. (2010). Improving your data transformations: applying the box-cox transformation. *Practical Assessment, Research, & Evaluation, 15*(12), 2-9.
- Ozgungor, S., & Guthrie, J. T. (2004). Interactions among elaborative interrogation, knowledge, and interest in the process of constructing knowledge from text. *Journal of Educational Psychology, 96*(3), 437-443.
- Pask, G., (1976). Styles and strategies of learning. *British Journal of Educational Psychology, 46*(2), 128-148.
- Payne, S. J., & Reader, W. R. (2006). Constructing structure maps of multiple on-line texts. *International Journal of Human-Computer Studies, 64*(5), 461-474.
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research: explanation and prediction*. New York: Holt, Rinehart, & Winston.
- Perfetti, C. A., Britt, M. A., Rouet, J. F., Georgi, M. C., & Mason, R. A. (1994). How students use texts to learn and reason about historical uncertainty. In Carretero, M. & Voss, J. F., (Eds.), *Cognitive and instructional processes in history and the social sciences* (pp. 257-283). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Perfetti, C. A., Rouet, J. F., & Britt, M. A., (1999). Toward a theory of document representation. In van Oostendorp, H., & Goldman, S. R., (Eds.), *The construction of mental representations during reading* (pp. 99-122). Mahwah, NJ: Lawrence Erlbaum Associates.

- Perry, W. (1981). Cognitive and ethical growth: The making of meaning. In P. A. Altbach, K. Arnold, & I. C. King (Eds.), *College student development and academic life: Psychological, intellectual, social, and moral issues* (pp. 48-88). New York: Routledge.
- Perry, W. G. (1970). *Forms of intellectual and ethical development*. New York: Holt, Rinehart and Winston.
- Petty, R. E., & Brinol, P. (2010). Attitude change. In R. F. Baumeister & E. J. Finkel (Eds.), *Advanced Social Psychology: The State of the Science*, (pp. 217-259). New York: Oxford University Press.
- Petty, R. E., Wegener, D. T., & Fabrigar, L. R. (1997). Attitudes and attitude change. *Annual Review of Psychology*, 48(1), 609-647.
- Potelle, H., & Rouet, J. F. (2003). Effects of content representation and readers' prior knowledge on the comprehension of hypertext. *International Journal of Human-Computer Studies*, 58(3), 327-345.
- Pratkanis, A. R. (1989). The cognitive representation of attitudes. In A. R. Pratkanis, S. J. Beckler, & A. G. Greenwald (Eds.) *Attitude Structure and Function*. Hillsdale, NJ: Lawrence Erlbaum (pp. 71-98).
- Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Lawrence Erlbaum.
- Purcell, K., Heaps, A., Buchanan, J., & Friedrich, L. (2013). How teachers are using technology at home and in their classrooms. *Pew Research Center's Internet & American Life Project*. Retrieved from:

<http://www.pewinternet.org/2013/02/28/how-teachers-are-using-technology-at-home-and-in-their-classrooms/>

- Purcell, K., Rainie, L., Heaps, A., Buchanan, J., Friedrich, L., Jacklin, A., ... Zickuhr, K. (2012). *How teens do research in the digital world*. Pew Internet & American Life Project. Retrieved from http://www.pewinternet.org/~media/Files/Reports/2012/PIP_TeacherSurveyReportWithMethodology110112.pdf
- Purdy, J. P. (2012). Why first-year college students select online research resources as their favorite. *First Monday*, 17(9). Retrieved from: <http://pear.acc.uic.edu/ojs/index.php/fm/article/view/4088/3289>
- Ratcliffe, J. F. (1968). The effect on the t distribution of non-normality in the sampled population. *Applied Statistics*, 17(1), 42-48.
- Reader, W. R., & Payne, S. J. (2007). Allocating time across multiple texts: sampling and satisficing. *Human-Computer Interaction*, 22(3), 263-298.
- Renninger, A., Hidi, S., & Krapp, A. (Eds.). (2014). *The role of interest in learning and development*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Riding, R. J., & Al-Sanabani, S. (1998). The effect of cognitive style, age, gender and structure on the recall of prose passages. *International Journal of Educational Research*, 29(3), 173-185.
- Rothenberg, D. (1998). How the web destroys student research papers. *Education Digest*, 63(6), 59-61.

- Rouet, J. F. (2003). What was I looking for?: the influence of task specificity and prior knowledge on students' search strategies in hypertext. *Interacting with Computers, 15*(3), 409-428.
- Rouet, J. F. (2006). *The skills of document use: From text comprehension to Web-based learning*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Rouet, J. F., & Britt, M. A. (2011). Relevance processing in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19-52). Greenwich, CN: Information Age Publishing, Inc.
- Rouet, J. F., Britt, M. A., Mason, R. A., & Perfetti, C. A. (1996). Using multiple sources of evidence to reason about history. *Journal of Educational Psychology, 88*(3), 478-493.
- 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., Levonen, J. J., Dillon, A., & Spiro, R. J. (Eds.) (1996) *Hypertext and Cognition*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Rouet, J. F., Ros, C., Goumi, A., Macedo-Rouet, M., & Dinet, J. (2011). The influence of surface and deep cues on primary and secondary school students' assessment of relevance in Web menus. *Learning and Instruction, 21*(2), 205-219.
- Rouet, J. F., Vidal-Abarca, E., Erbou, A. B., & Millogo, V. (2001). Effects of information search tasks on the comprehension of instructional text. *Discourse Processes, 31*(2), 163-186.

- Salmerón, L., Gil, L., Bråten, I., & Strømsø, H. (2010). Comprehension effects of signalling relationships between documents in search engines. *Computers in Human Behavior, 26*(3), 419-426.
- Salmerón, L., Kintsch, W., & Cañas, J. J. (2006). Reading strategies and prior knowledge in learning from hypertext. *Memory & Cognition, 34*(5), 1157-1171.
- Salmeron, L., Kintsch, W., & Canas, J. (2006). Coherence or interest as basis for improving hypertext comprehension. *Information Design Journal, 14*(1), 45-55.
- Salmerón, L., Kintsch, W., & Kintsch, E. (2010). Self-regulation and link selection strategies in hypertext. *Discourse Processes, 47*(3), 175-211.
- Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition, 34*(2), 344-355.
- Sawilowsky, S. S., & Blair, R. C. (1992). A more realistic look at the robustness and Type II error properties of the t-test to departures from population normality. *Psychological Bulletin, 111*(2), 352-360.
- Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist, 26*(3/4), 299-323.
- Schiefele, U. (1999). Interest and learning from text. *Scientific Studies of Reading, 3*(3), 257-279.
- Schiefele, U., & Krapp, A. (1996). Topic interest and free recall of expository text. *Learning and Individual Differences, 8*(2), 141-160.
- Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust?: reinvestigating the robustness of ANOVA against violations of the normal

- distribution assumption. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 6(4), 147-151.
- Schrader, P. G., & Lawless, K. A. (2004). The knowledge, attitudes, and behaviors approach how to evaluate performance and learning in complex environments. *Performance Improvement*, 43(9), 8-15.
- Shah, P., Mayer, R. E., & Hegarty, M. (1999). Graphs as aids to knowledge construction: signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology*, 91(4), 690-702.
- Shapiro, A. M. (1999). The relationship between prior knowledge and interactive overviews during hypermedia-aided learning. *Journal of Educational Computing Research*, 20(2), 143-167.
- Shapiro, A., & Niederhauser, D. (2004). Learning from hypertext: Research issues and findings. In: D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2nd ed, pp. 605-620). Mahwah, NJ: Erlbaum.
- Shuart-Faris, N., & Bloome, D. (Eds.). (2004). *Uses of intertextuality in classroom and educational research*. Greenwich, CT: Information Age Publishing.
- Smith, A. G. (1997). Testing the surf: criteria for evaluating Internet information resources. *Public Access-Computer Systems Review*, 8(3), 5-23.
- Snider, E. A., & Faris, D. M. (2011). The Arab spring: US democracy promotion in Egypt. *Middle East Policy*, 18(3), 49-62.
- Sparks, P. C. (2013). *Do task instructions influence readers' topic beliefs, topic belief justifications, and task interest? A mixed methods study*. Retrieved from:

<http://researcharchive.vuw.ac.nz/xmlui/bitstream/handle/10063/2844/thesis.pdf?sequence=2>

- Stadtler, M., & Bromme, R. (2007). Dealing with multiple documents on the WWW: the role of metacognition in the formation of documents models. *International Journal of Computer-Supported Collaborative Learning*, 2(2-3), 191-210.
- Stadtler, M., & Bromme, R. (2008). Effects of the metacognitive computer-tool “met. a. ware” on the web search of laypersons. *Computers in Human Behavior*, 24(3), 716-737.
- Stadtler, M., Scharrer, L., Brummernhenrich, B., & Bromme, R. (2013). Dealing with uncertainty: readers' memory for and use of conflicting information from science texts as function of presentation format and source expertise. *Cognition and Instruction*, 31(2), 130-150.
- 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.
- Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences*. New York: Routledge.
- Strømsø, H. I., & Bråten, I. (2002). Norwegian law students' use of multiple sources while reading expository texts. *Reading Research Quarterly*, 37(2), 208-227.
- Strømsø, H. I., Bråten, I., & Britt, M. A. (2010). Reading multiple texts about climate change: the relationship between memory for sources and text comprehension. *Learning and Instruction*, 20(3), 192-204.

- Strømsø, H. I., Bråten, I., & Britt, M. A. (2011). Do students' beliefs about knowledge and knowing predict their judgment of texts' trustworthiness? *Educational Psychology: An International Journal of Experimental Educational Psychology*, 31(2), 177-206.
- 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.
- Strømsø, H. I., Bråten, I., & Samuelstuen, M. S. (2008). Dimensions of topic-specific epistemological beliefs as predictors of multiple text understanding. *Learning and Instruction*, 18(6), 513-527.
- Strømsø, H. I., Bråten, I., Britt, M. A., & Ferguson, L. E. (2013). Spontaneous sourcing among students reading multiple documents. *Cognition and Instruction*, 31(2), 176-203.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4(4), 295-312.
- Symons, S., Pressley, M., 1993. Prior knowledge affects text search success and extraction of information. *Reading Research Quarterly*, 28(3), 250–261.
- Tabachnick, B. G., & Fidell, L. S. (2001). Using multivariate statistic. (4th ed.) Boston, MA: Allyn & Bacon/Pearson Education.
- Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50(3), 755-769.
- Taber, C. S., Cann, D., & Kucsova, S. (2009). The motivated processing of political arguments. *Political Behavior*, 31(2), 137-155.

University of Maryland Office of Institutional Research Planning and Assessment

Retrieved from:

https://irpa.umd.edu/CampusCounts/Enrollments/stuprofile_allug.pdf

van Oostendorp, H., & De Mul, S. (Eds.). (1996). *Cognitive aspects of electronic text processing*. Norwood, New Jersey: Greenwood Publishing Group.

van Strien, J. L., Brand-Gruwel, S., & Boshuizen, H. (2014). Dealing with conflicting information from multiple nonlinear texts: Effects of prior attitudes. *Computers in Human Behavior*, 32, 101-111.

van Strien, J., Brand-Gruwel, S., & Boshuizen, E. (2012). Do prior attitudes influence epistemic cognition while reading conflicting information?. Poster presented at the Sig 2 EARLI conference, Grenoble, France. Retrieved from:

<http://dspace.ou.nl/bitstream/1820/4390/3/33%20VAN%20STRIEN.pdf>

VanSledright, B. A. (2002). Fifth graders investigating history in the classroom: results from a researcher-practitioner design experiment. *The Elementary School Journal*, 131-160.

Wade, S. E., Buxton, W. M., & Kelly, M. (1999). Using think-alouds to examine reader-text interest. *Reading Research Quarterly*, 34(2), 194-216.

Wade, S. E., Buxton, W. M., & Kelly, M. (1999). Using think-alouds to examine reader-text interest. *Reading Research Quarterly*, 34(2), 194-216.

Wallace, R. M., Kupperman, J., Krajcik, J., & Soloway, E. (2000). Science on the Web: students online in a sixth-grade classroom. *The Journal of the Learning Sciences*, 9(1), 75-104.

- Walraven, A., Brand-Gruwel, S., & Boshuizen, H. (2009). How students evaluate information and sources when searching the World Wide Web for information. *Computers and Education*, 52(1), 234-246.
- Wang, Y. M., & Artero, M. (2005). Caught in the Web: university student use of Web resources. *Educational Media International*, 42(1), 71-82.
- Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997): Overall job satisfaction: how good are single-item measures? *Journal of Applied Psychology*, 82(2), 247–252.
- Weinberg, S. L., & Abramowitz, S. K. (2008). *Statistics using SPSS: An integrative approach*. New York: Cambridge University Press.
- Weinreich, H., Obendorf, H., Herder, E., & Mayer, M. (2008). Not quite the average: an empirical study of Web use. *ACM Transactions on the Web*, 2(1), 1-31.
- Weinstein. C E ., Palmer. D. R., & Schulte. A C. (1987) *Learning and study strategies inventory*. Clearwater, FL: H & H Publishing.
- Wenger, M. J., & Payne, D. G. (1996). Comprehension and retention of nonlinear text: considerations of working memory and material-appropriate processing. *The American Journal of Psychology*, 109(9), 93-130.
- Wigfield, A., & Guthrie, J. T. (2000). Engagement and motivation in reading. In M. Kamil, R. Barr, P. Mosenthal, & D. Pearson (Eds.), *Handbook of Reading Research III*, (pp. 303-422). Mahwah, NJ: Lawrence Erlbaum.
- 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(2), 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.
- Wilkinson, S. C., Reader, W., & Payne, S. J. (2012). Adaptive browsing: sensitivity to time pressure and task difficulty. *International Journal of Human-Computer Studies*, *70*(1), 14-25.
- Wilson, D., & Sperber, D. (2002). Relevance theory. In: G. Ward & L. Horn (Eds.), *Handbook of Pragmatics*, (pp. 607-632). Oxford, UK: Blackwell Science.
- Wineburg, S. 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*(1), 73-87. doi: [10.1037/0022-0663.83.1.73](https://doi.org/10.1037/0022-0663.83.1.73)
- Wolfe, C. R., & Britt, M. A. (2008). The locus of the myside bias in written argumentation. *Thinking & Reasoning*, *14*(1), 1-27.
- Wolfe, C. R., Britt, M. A., & Butler, J. A. (2009). Argumentation schema and the myside bias in written argumentation. *Written Communication*, *26*(2), 183-209.
- Wolfe, M. B., & Goldman, S. R. (2005). Relations between adolescents' text processing and reasoning. *Cognition and Instruction*, *23*(4), 467-502.
- Wood, W., Rhodes, N., & Biek, M. (1995). Working knowledge and attitude strength: An information-processing analysis. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 189-202) New York: Psychology Press.